LinDoA for Hearing Aids

Direction of Arrival Estimation and Beamforming Using a Taylor Series Expansion

Clas Veibäck

LINKÖPING UNIVERSITY

About me

PhD in Automatic Control

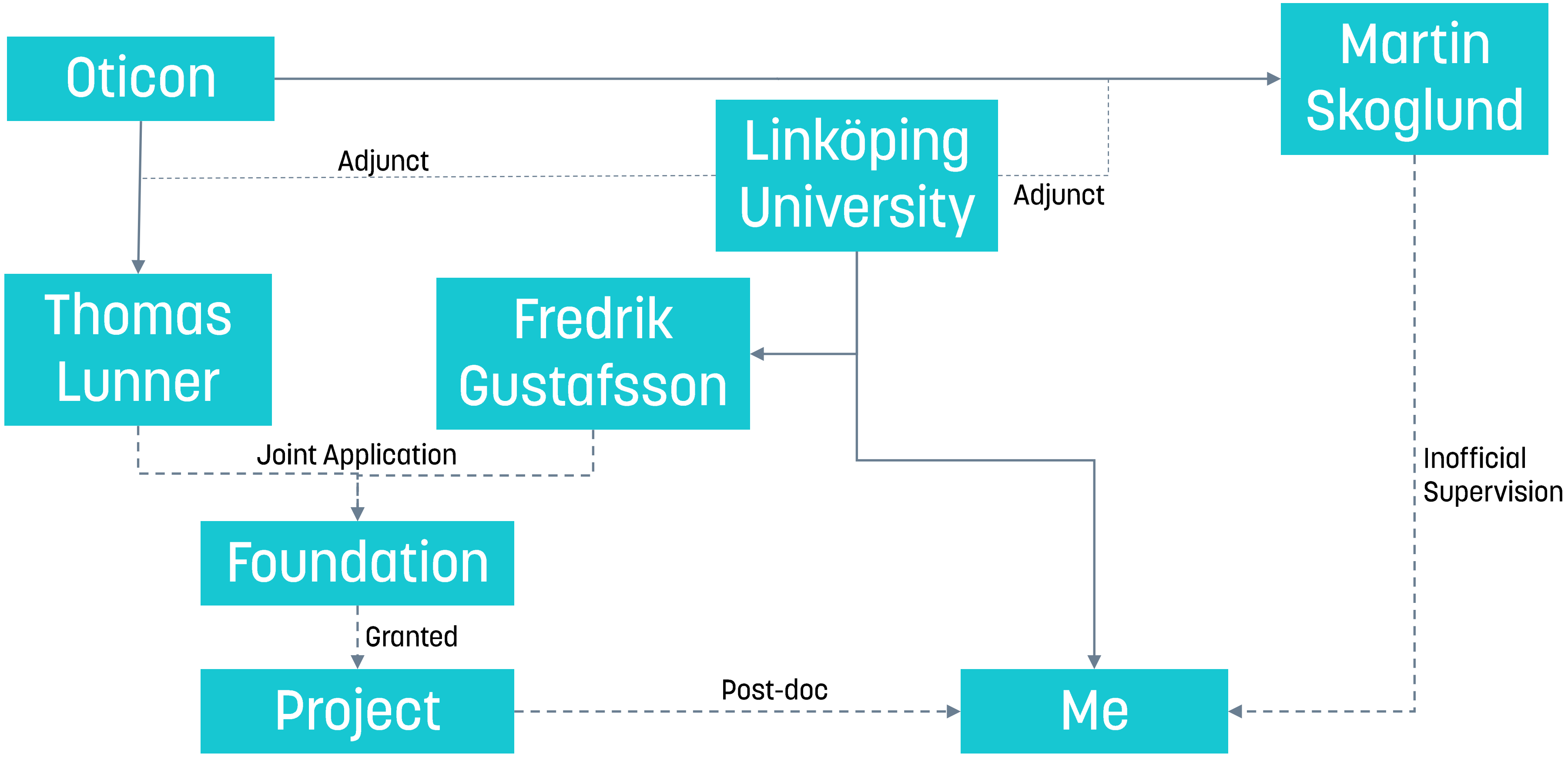
Target tracking

Post-doc

William Demant foundation

Parental leave

How Things Connect



Project Description

Target tracking

Direction of arrival (DoA) estimation

Beamforming

Local Taylor-series expansion

Adapt and verify

Develop prototype

Experiments

Potential Advantages

General geometries

Near and far field

Ultra-slow sampling

Small dimensions

Broadband signals

Parallelization

Model

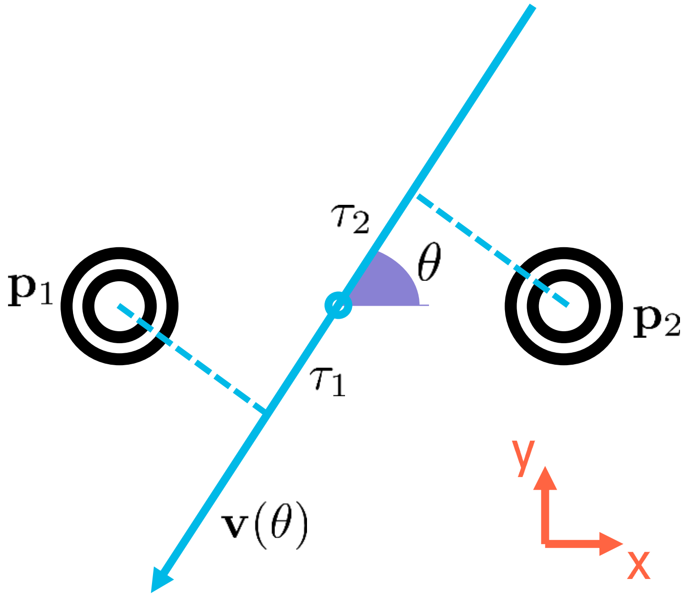
Delays are computed geometrically

$$\tau_m = \mathbf{p}_m^T \mathbf{v}(\theta)$$

$$\mathbf{v}(\theta) = \frac{1}{c} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

Azimuth and elevation

Sampled signal $y_k^m = s(t_k + \tau_m) + e_k^m$



Delay and Sum

1. Compute delays
2. Align
3. Average

Amplifies given direction

Attenuates other directions

Issue: Continuous time shift on discrete signal?

Solution: Interpolation

Frequency domain

Frequency Domain

$$\mathbf{Y}(\omega) = \begin{bmatrix} Y_1(\omega) \\ \vdots \\ Y_M(\omega) \end{bmatrix} = \begin{bmatrix} \exp(i\omega\tau_1) \\ \vdots \\ \exp(i\omega\tau_M) \end{bmatrix} S(\omega) + \begin{bmatrix} E_1(\omega) \\ \vdots \\ E_M(\omega) \end{bmatrix} = \mathbf{a}(\theta, \omega)S(\omega) + \mathbf{E}(\omega)$$

DS corresponds to $\hat{S}(\omega) = \frac{1}{M}\mathbf{a}^H(\theta, \omega)\mathbf{Y}(\omega)$

Minimum-Variance Distortionless Response

Find filter $\hat{S}(\omega) = \mathbf{w}^H \mathbf{Y}(\omega) = \mathbf{w}^H \mathbf{a}(\theta, \omega) S(\omega) + \mathbf{w}^H \mathbf{E}(\omega)$

Optimization problem: $\arg\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{R}(\omega) \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{a}(\theta, \omega) = 1$

Solution: $\mathbf{w} = \frac{\mathbf{R}^{-1}(\omega)\mathbf{a}(\theta,\omega)}{\mathbf{a}^H(\theta,\omega)\mathbf{R}^{-1}(\omega)\mathbf{a}(\theta,\omega)}$

Issue: How to determine R?

LinDoA

Taylor expansion:
$$s(t+\tau) = s(t) + \tau s'(t) + \frac{\tau^2}{2} s''(t) + \dots$$
$$\approx s(t) + \tau s'(t) + \cdots + \frac{\tau^L}{L!} s^{(L)}(t) = \mathbf{h}(\tau)\mathbf{x}(t)$$

Decent approximation if:

- Small delays
- Small signal variation
- Large order

LinDoA

Model

$$\mathbf{y}_k = \mathbf{H}_{\boldsymbol{\tau}} \mathbf{x}_k + \mathbf{e}_k$$

where
$$\mathbf{H}_{\boldsymbol{\tau}} = \begin{pmatrix} 1 & \tau_1 & \dots & \tau_1^L \\ \vdots & \vdots & & \vdots \\ 1 & \tau_M & \dots & \tau_M^L \end{pmatrix}$$
and
$$\operatorname{cov}(\mathbf{e}_k) = \mathbf{R}$$

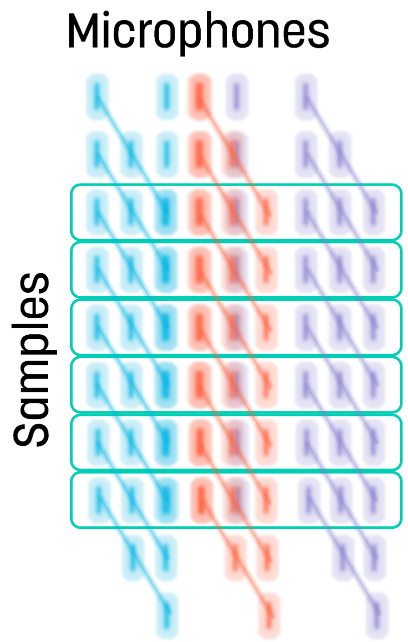
Generalized least squares

$$\hat{\mathbf{x}}_k = \arg\min_{\mathbf{x}} \ \|\mathbf{y}_k - \mathbf{H}_{\boldsymbol{\tau}} \mathbf{x}\|_{\mathbf{R}}^2 = (\mathbf{H}_{\boldsymbol{\tau}}^T \mathbf{R}^{-1} \mathbf{H}_{\boldsymbol{\tau}})^{-1} \mathbf{H}_{\boldsymbol{\tau}}^T \mathbf{R}^{-1} \mathbf{y}_k$$

LinDoA

- Correct model -> Interference attenuation
- Issue: Incrorrect model -> Poor interference attenuation
- Delay and Sum averages -> Interference attenuation
- Solution: DS in LinDoA form
- Solution: Correlation matrix

Extended LinDoA



Choice of Order

Signal variation

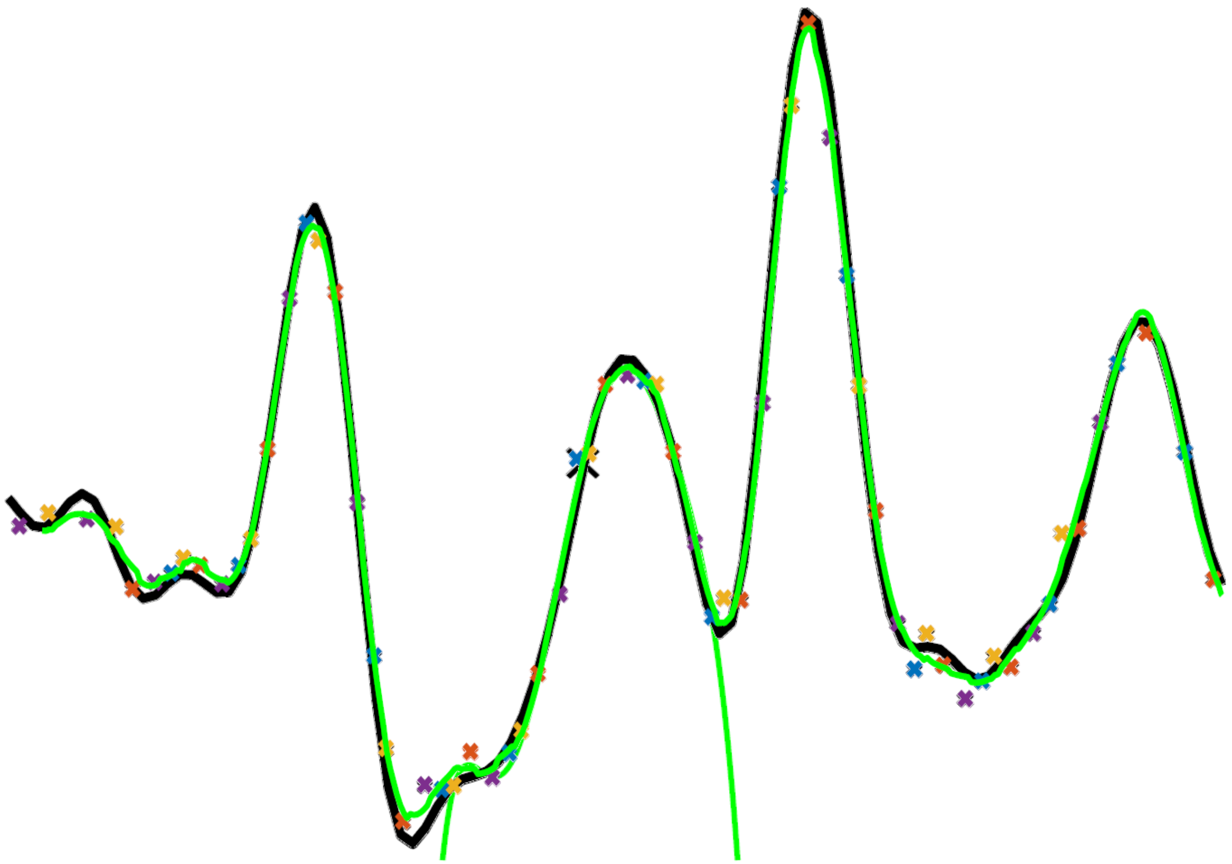
Number of samples

Time span

Computation increases with order

Solution: Samples close to center

Local Samples



Power or Error

LinDoA minimizes error $\hat{\theta} = \arg\min_\theta \sum_k \|\mathbf{y}_k - \mathbf{H}(\theta)\hat{\mathbf{x}}_k(\theta)\|_{\mathbf{R}}^2$

DS and MVDR minimize power $\hat{\theta} = \arg\min_\theta \sum_k \|\hat{s}_k(\theta)\|^2$

Minimum variance and average LinDoA can minimize power

Covariance Estimation

MVDR: One complex covariance matrix per frequency

Single rank update

Minimum variance LinDoA: One covariance matrix for sample extension

Full rank update

Voice activity detection

Filter LinDoA

State-space model

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{F}\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{H}_\tau \mathbf{x}_k + \mathbf{e}_k \end{aligned}$$

State transition matrix

$$\mathbf{F} = \begin{bmatrix} 1 & T & T^2 & \cdots & T^L \\ 0 & 1 & T & \cdots & T^{L-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & T \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

Smoothness: Order and process noise

Delayed Samples

State Estimation Using a Kalman Filter

Time update

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}\hat{\mathbf{x}}_{k|k}$$
$$\mathbf{P}_{k+1|k} = \mathbf{F}\mathbf{P}_{k|k}\mathbf{F}^T + \mathbf{Q}$$

Measurement update

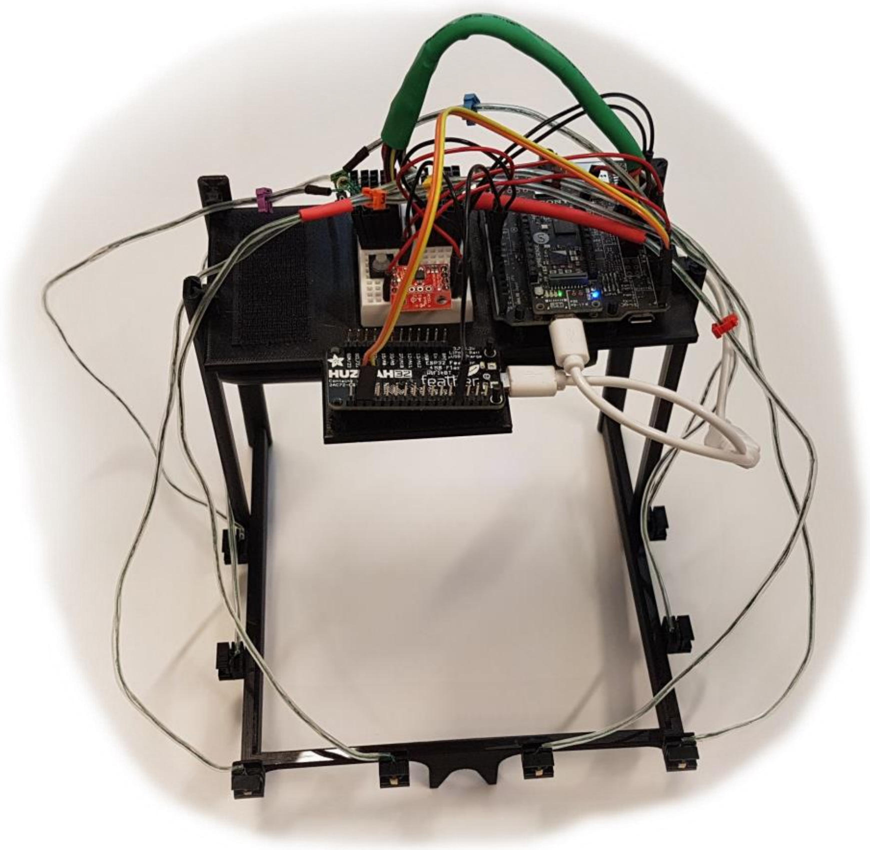
$$\mathbf{S}_k = \mathbf{H}_\tau \mathbf{P}_{k|k-1}\mathbf{H}_\tau^T + \mathbf{R}$$
$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_\tau^T \mathbf{S}_k^{-1}$$
$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_\tau \hat{\mathbf{x}}_{k|k-1})$$
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}\mathbf{S}_k\mathbf{K}_k^T$$

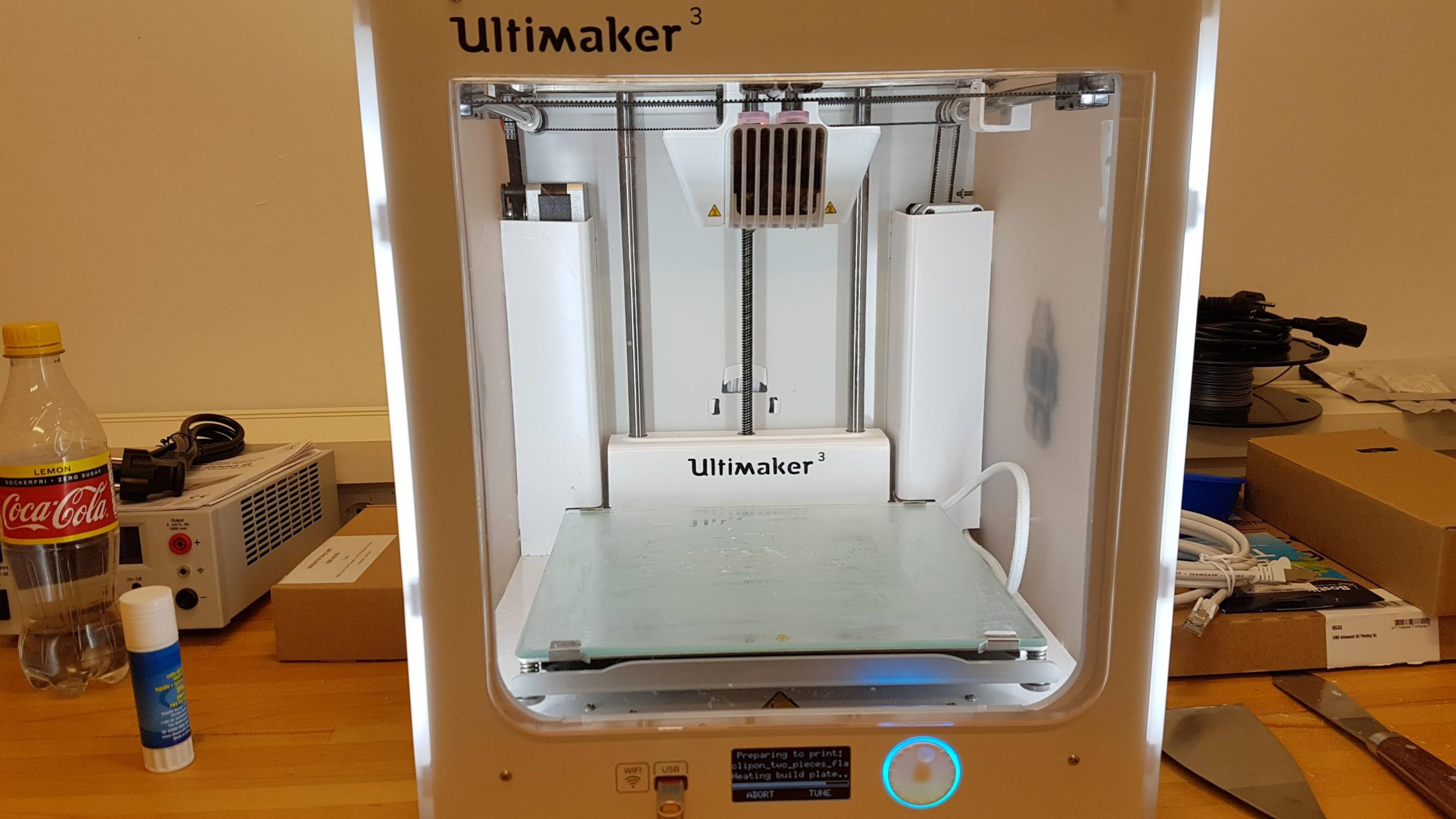
Stationary filter

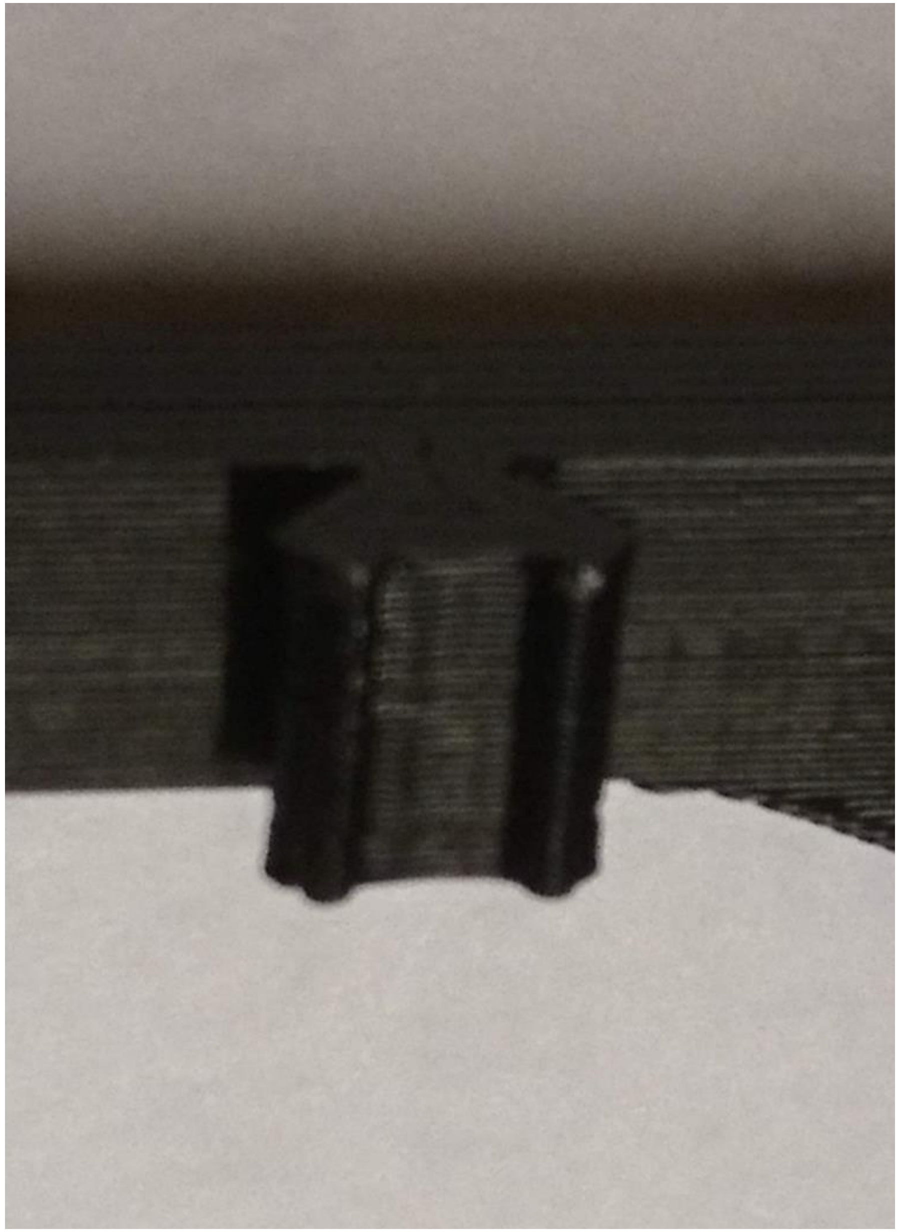
$$\hat{\mathbf{x}}_k = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{y}_k$$

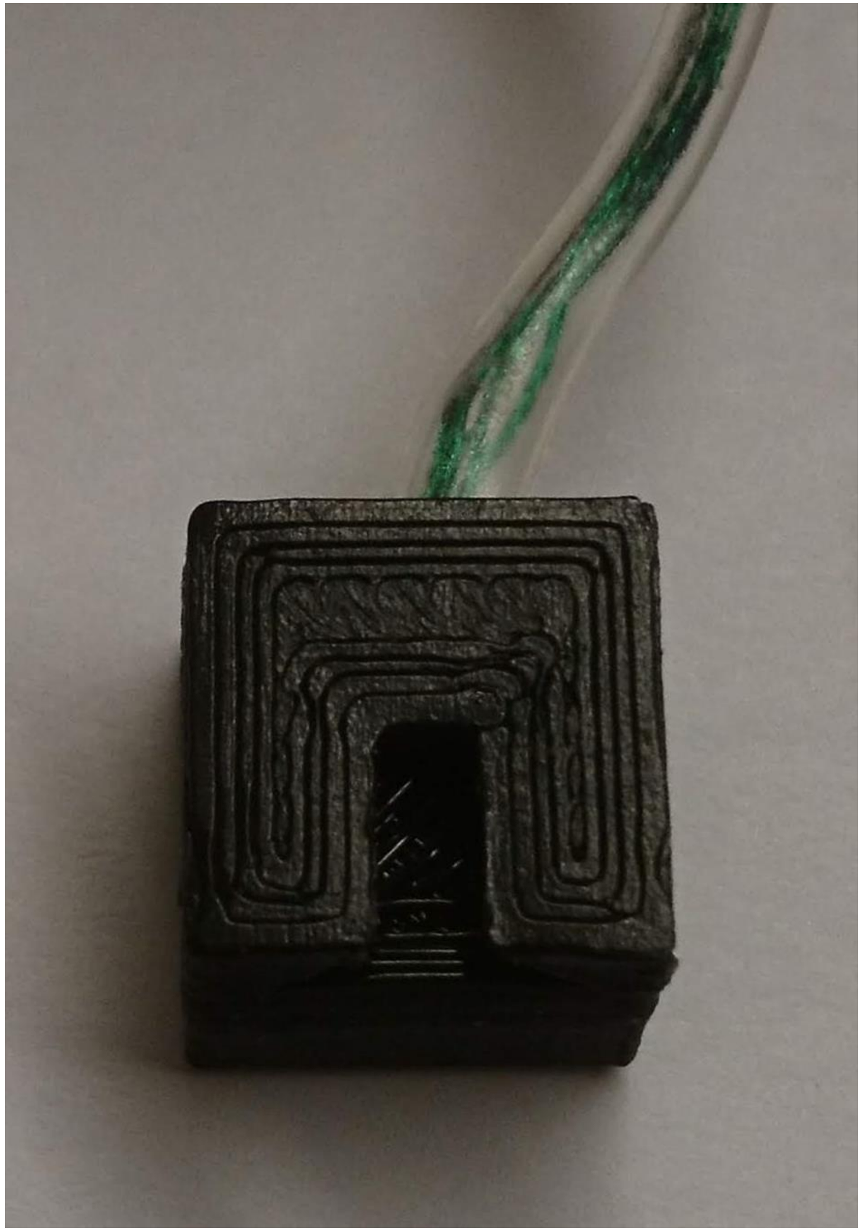
Prototype

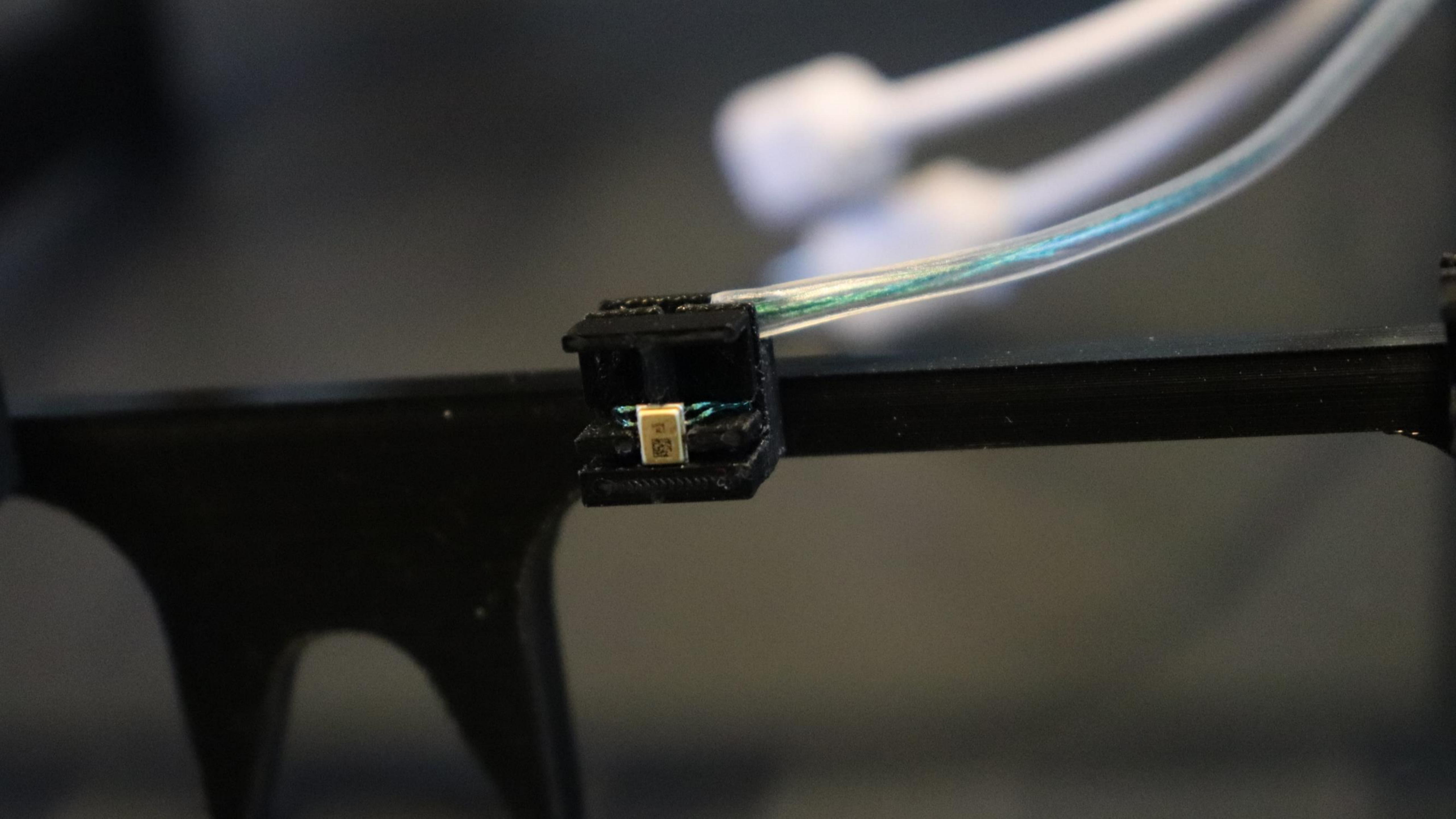
- Wearable ear-level array
- Glasses array geometry
- 3D printed
- Modular design
- Replace single parts
- Redesign array
- Sensor mounts

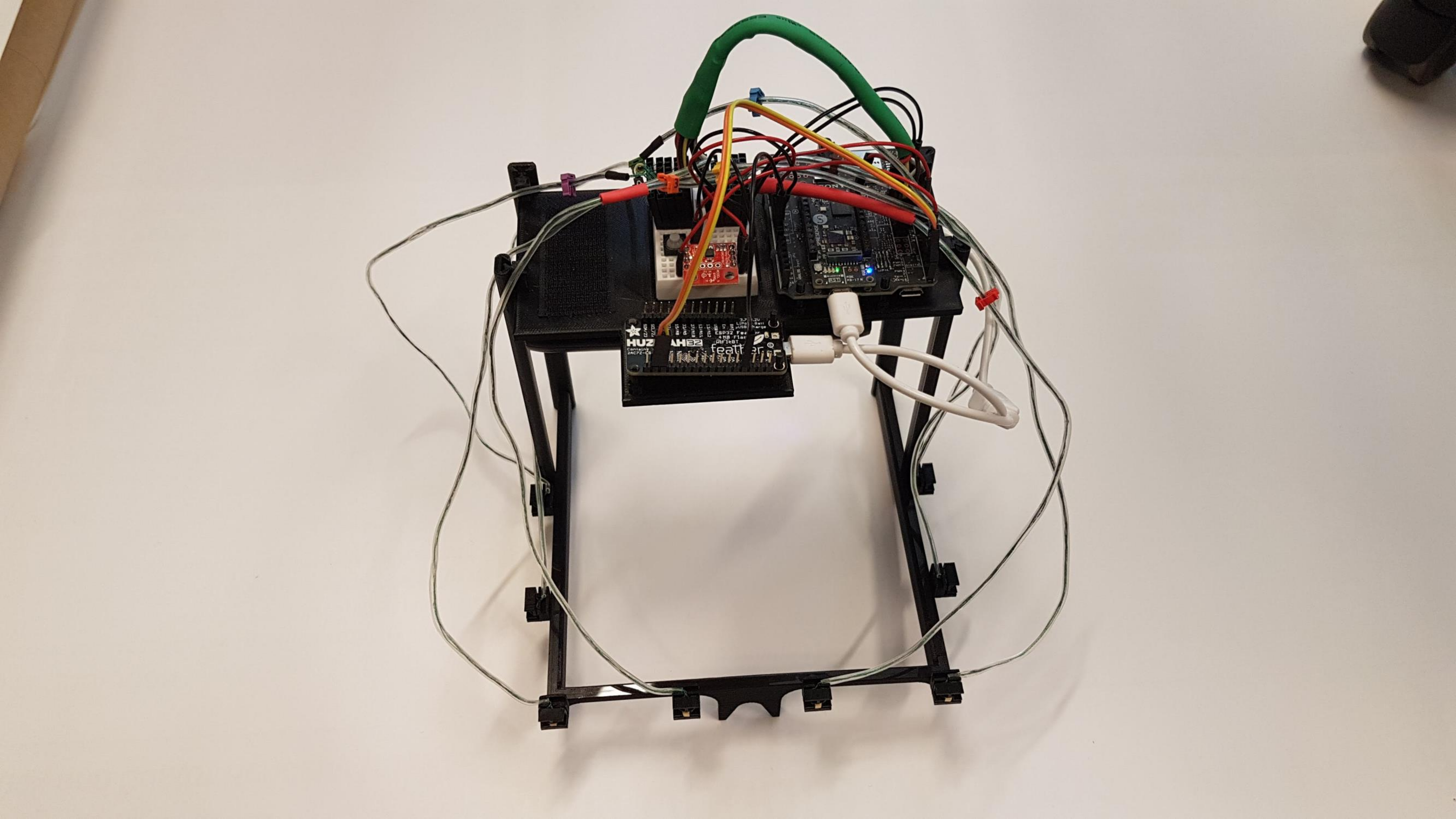


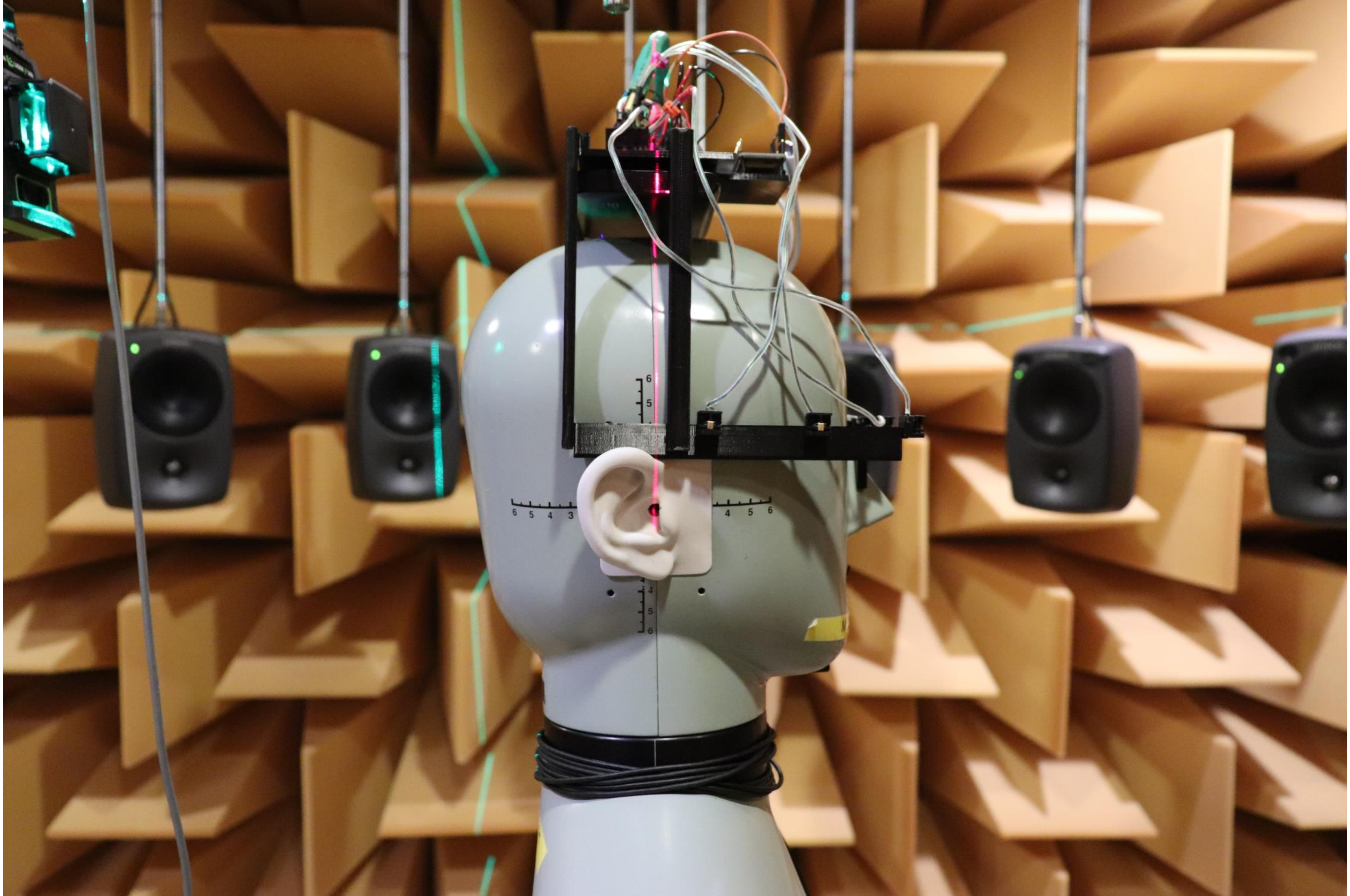

Ultimaker 3
Ultimaker 3
Preparing to print:
clipon_two_pieces_fla
Heating build plate..
ABORT
TUNE
WIFI
USB
LEMON
SOCKERFRI • ZERO SUGAR
Coca-Cola

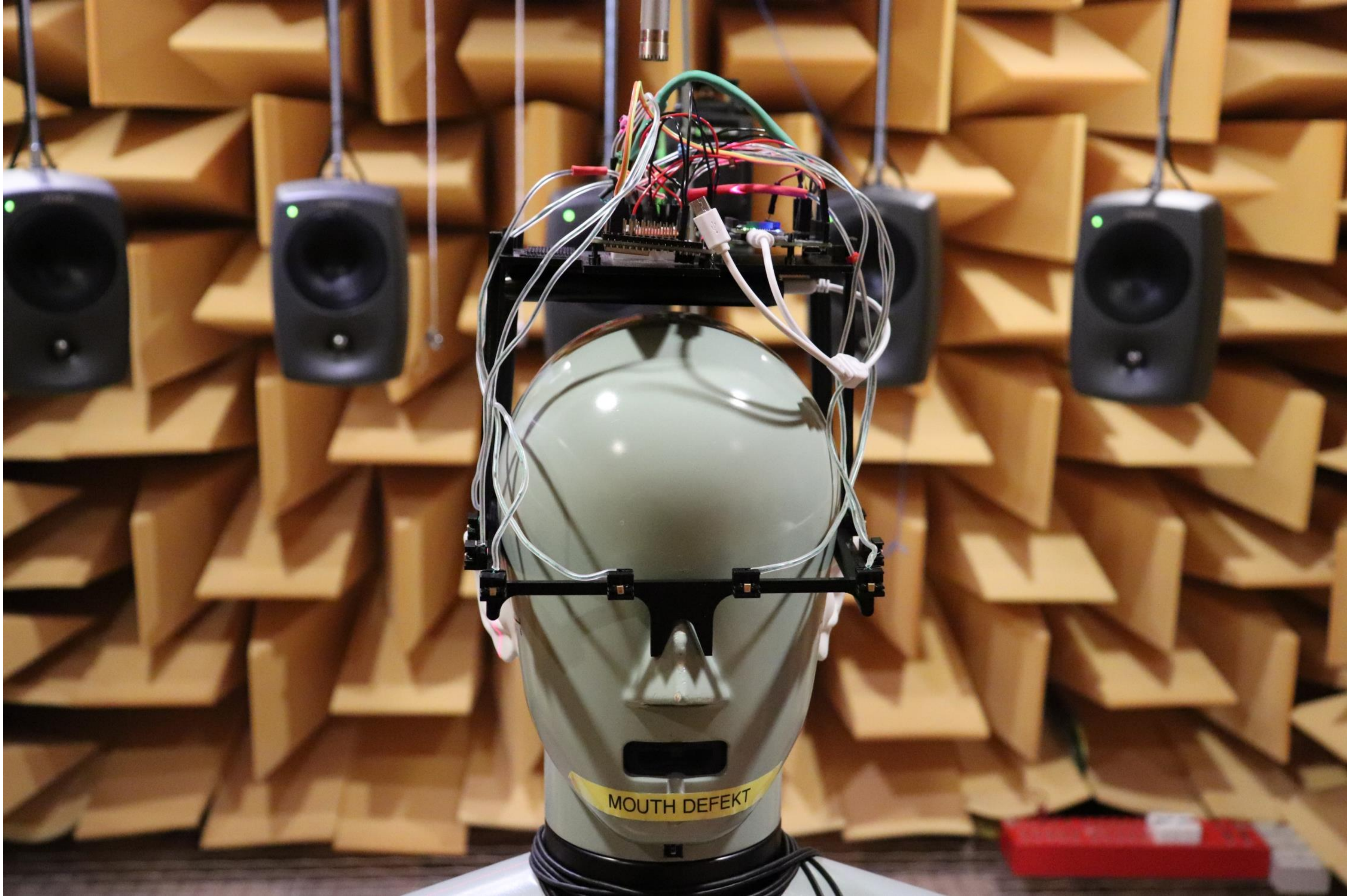











MOUTH DEFEKT

Electronics

DSP

8 microphones

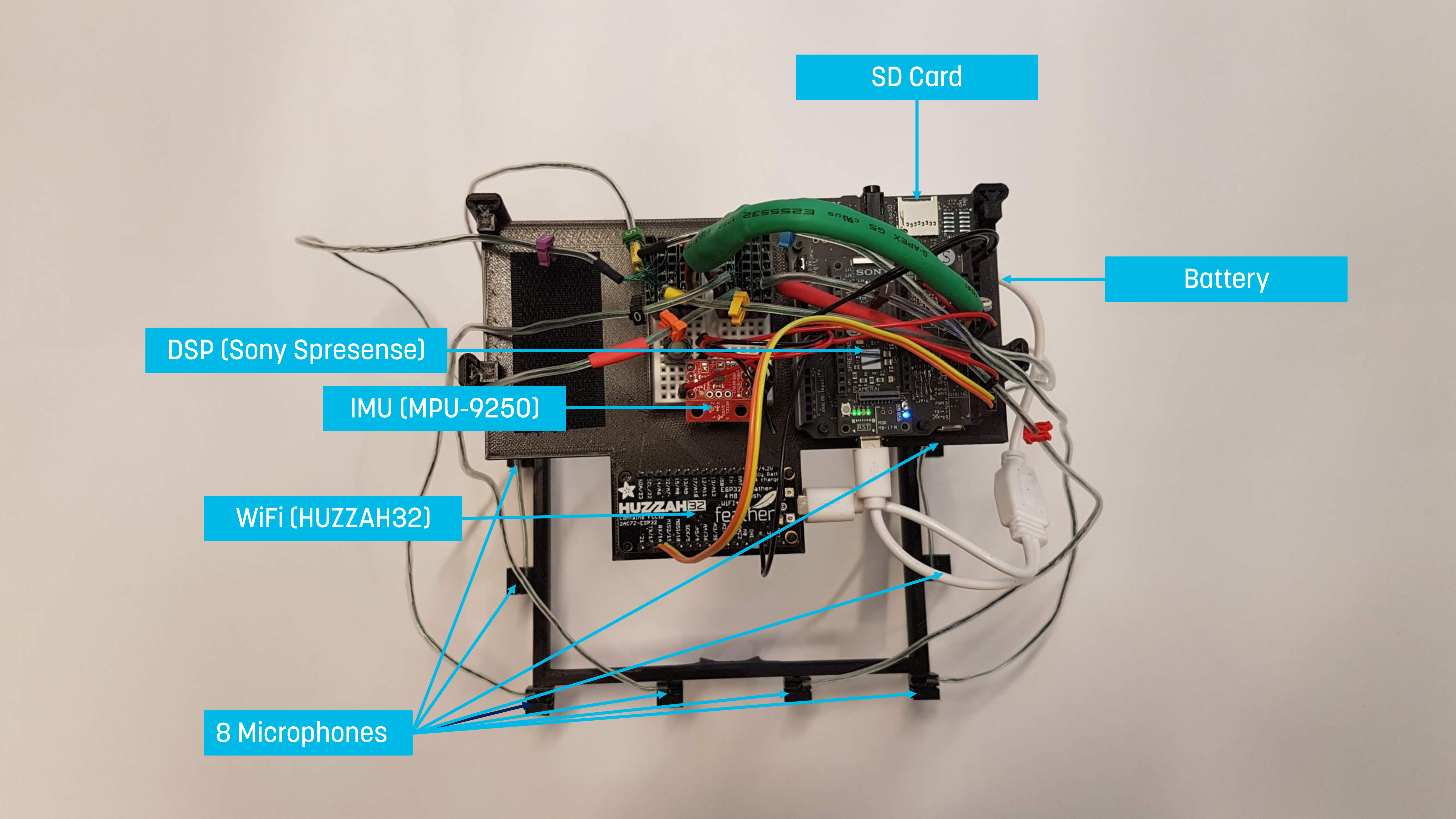
IMU

Synchronized

SD card

Battery

WiFi


SD Card
Battery
DSP (Sony Spresense)
IMU (MPU-9250)
WiFi (HUZZAH32)
8 Microphones

Experiments

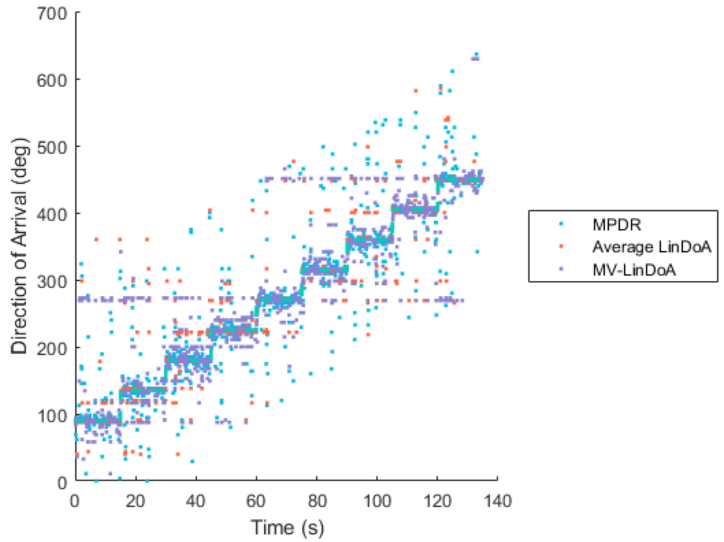
Reverberant environment

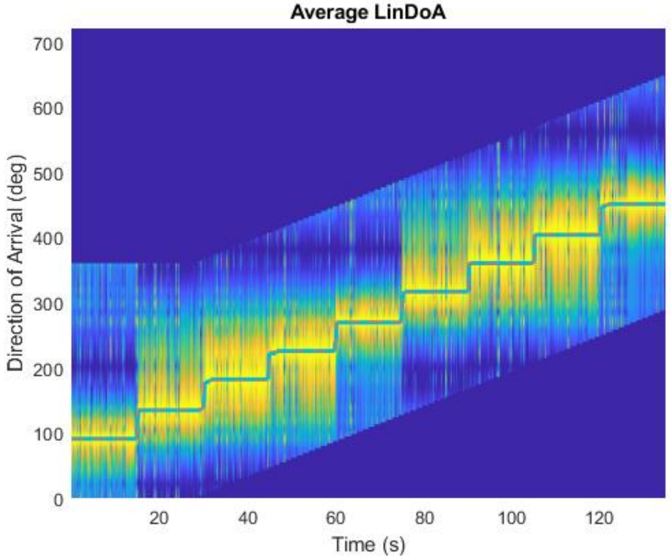
- Single and multiple sources
- Stationary and rotating array
- Simulation

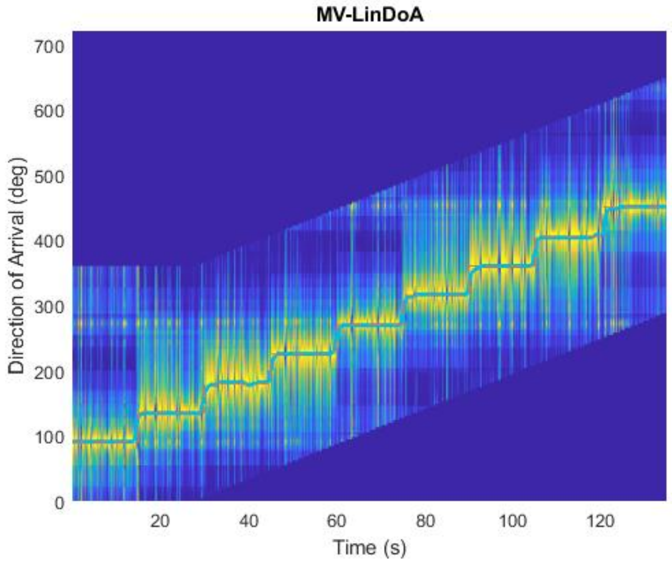
Anechoic environment

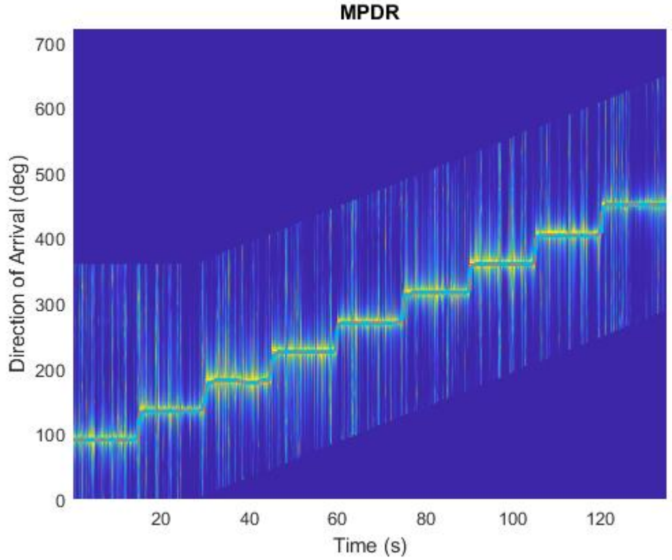
- Sweeps
- Dummy head

Simulation Rotating Monologue

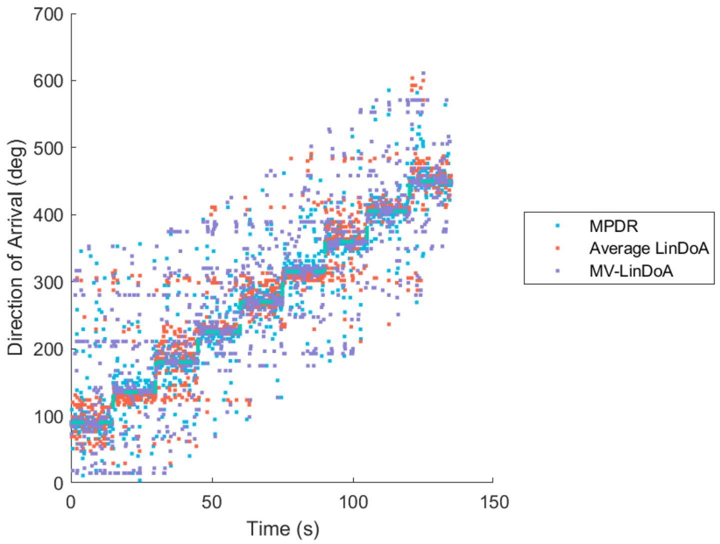


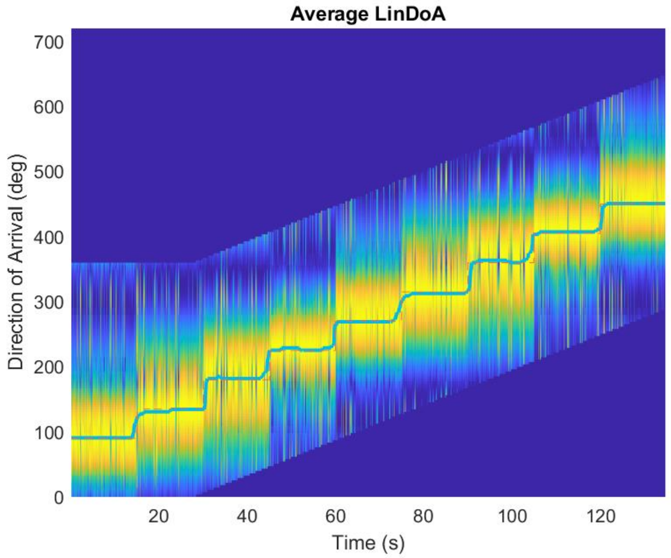


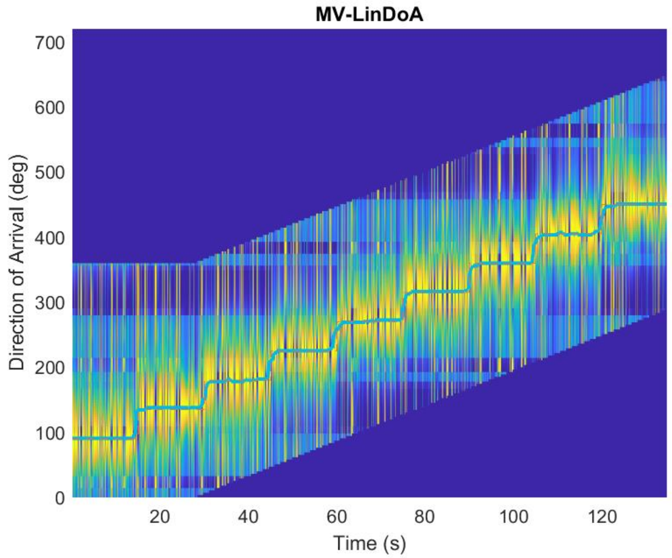


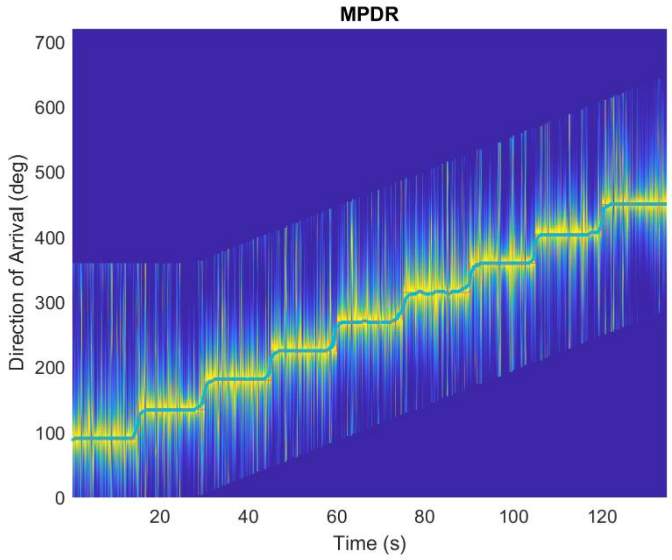


Simulation Rotating Small Monologue

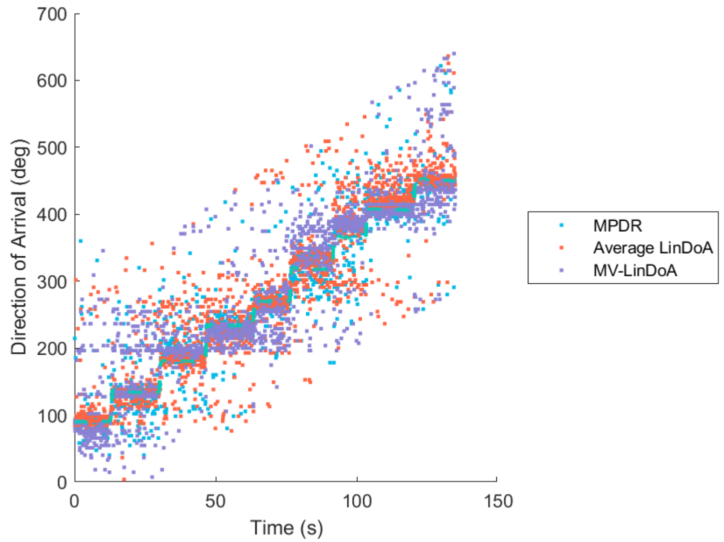


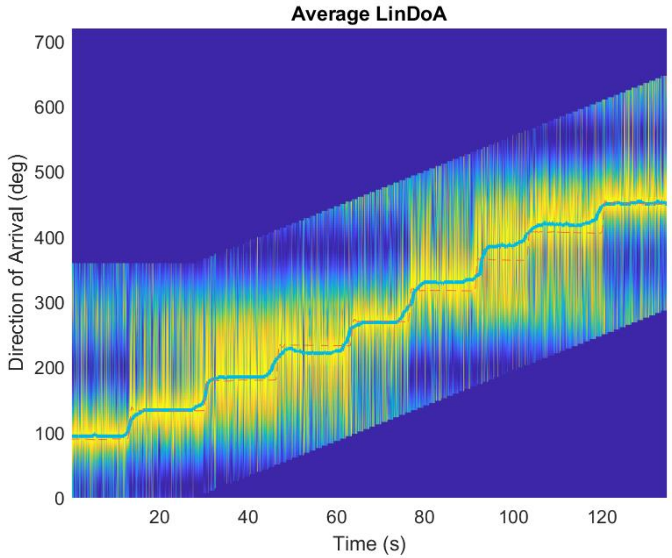


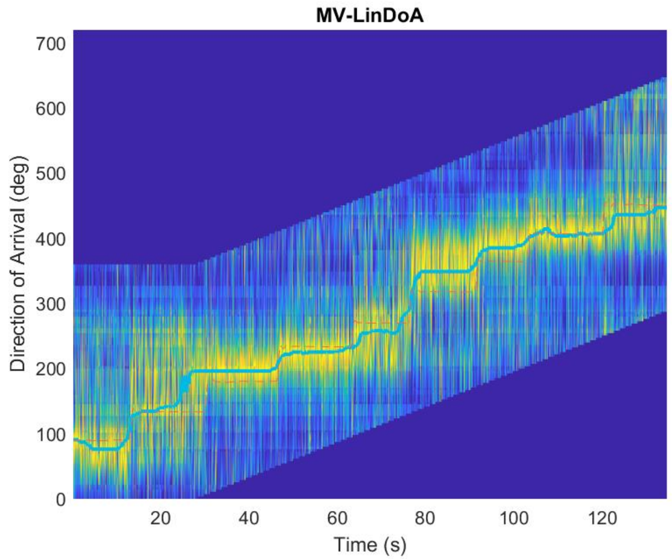


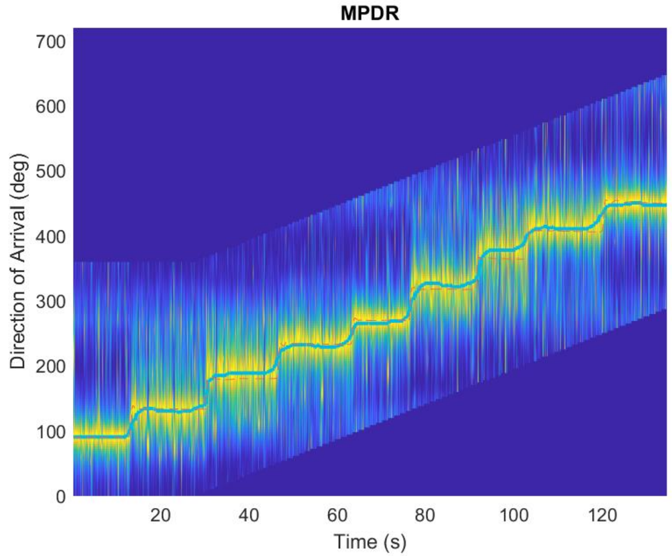


Hardware Rotating Monologue

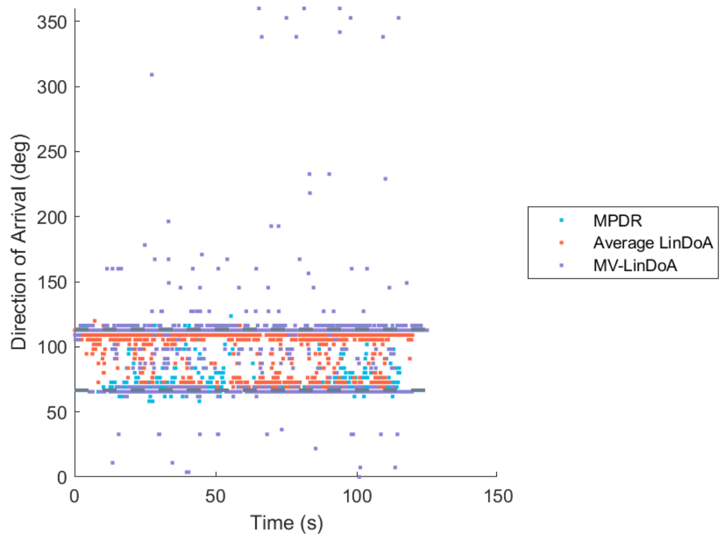


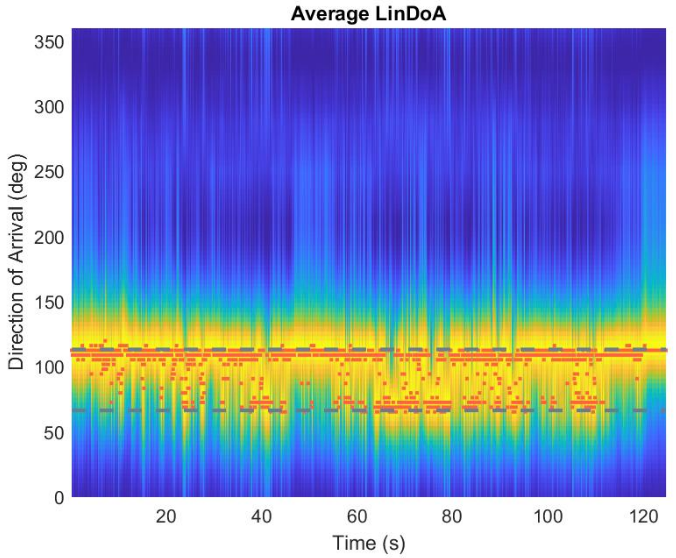


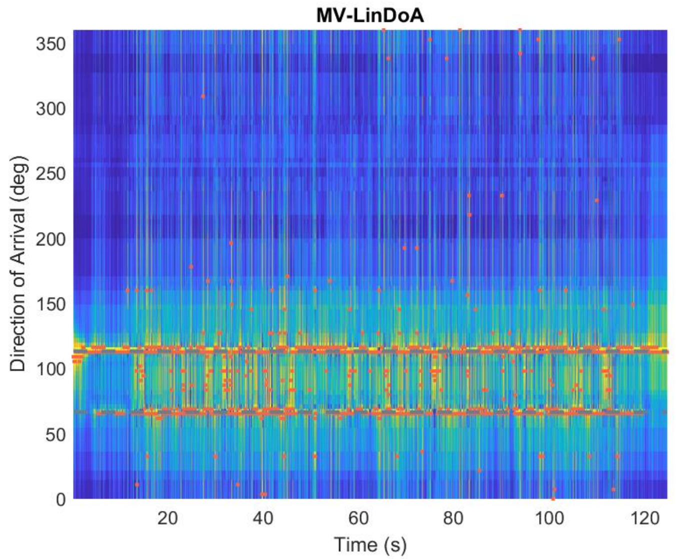


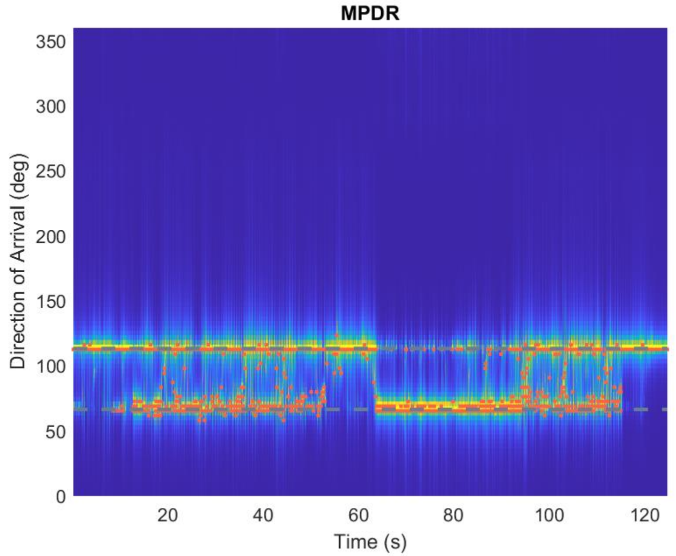


Simulation Still Music

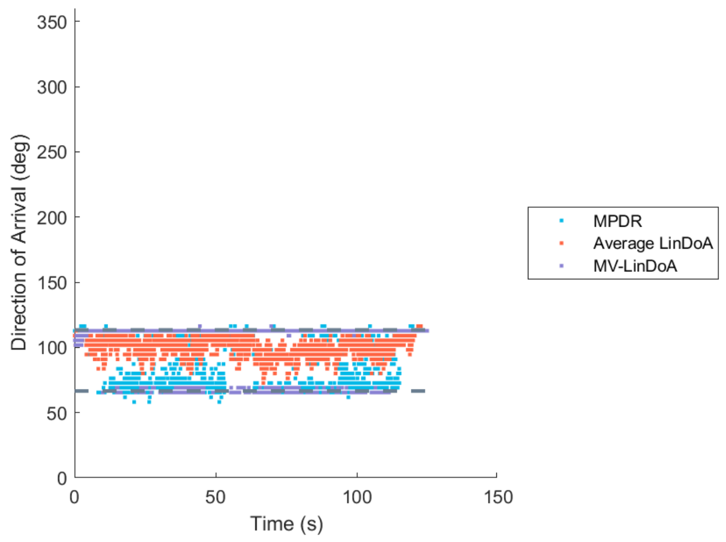


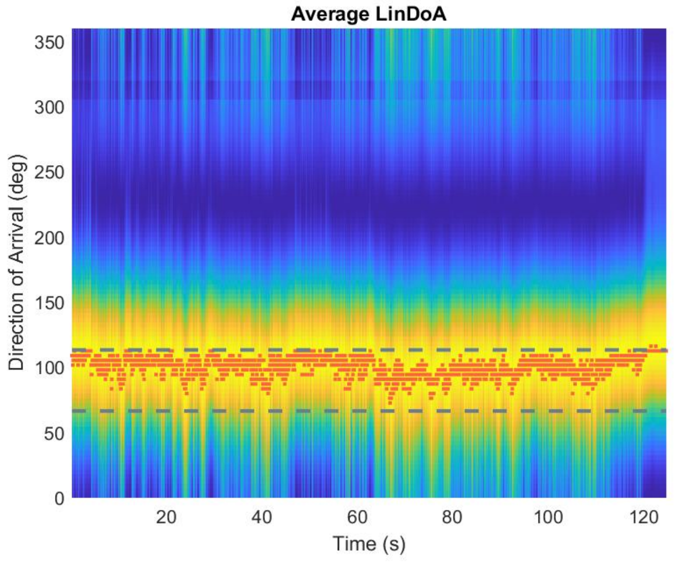


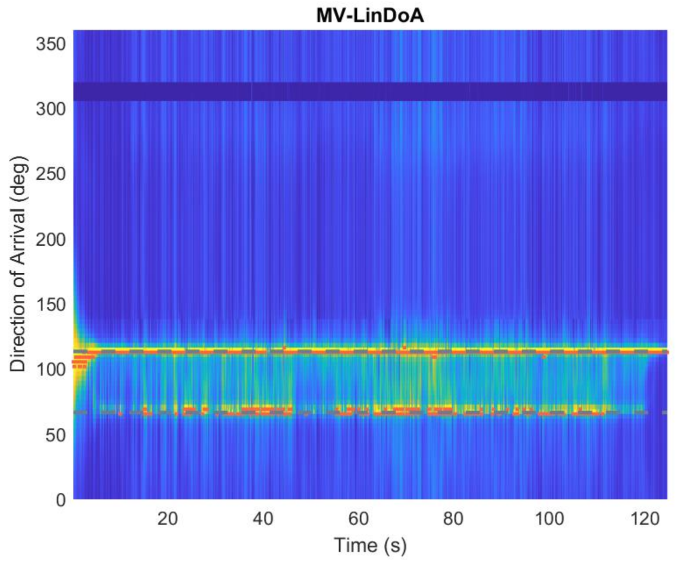


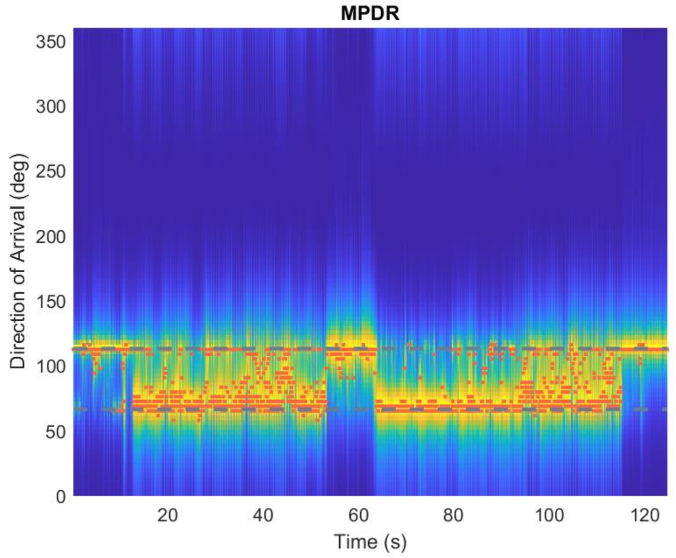


Simulation Still Small Music

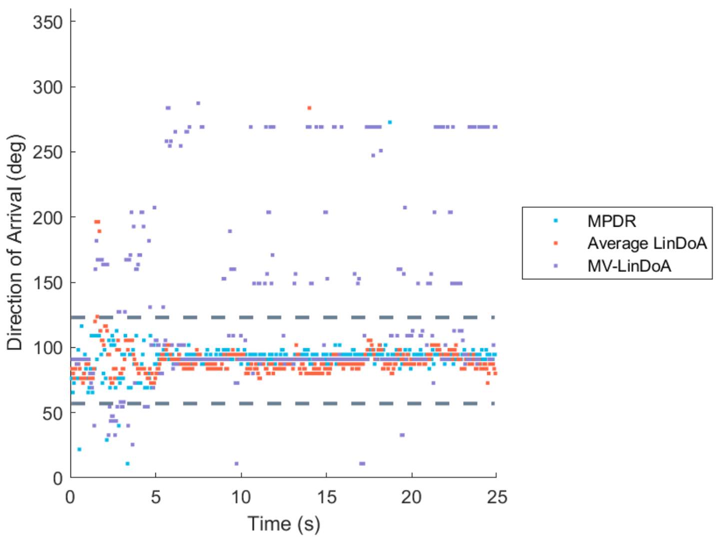


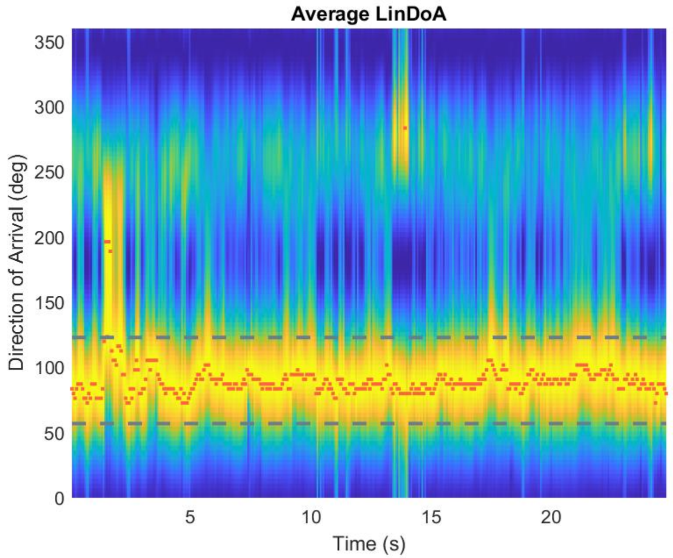


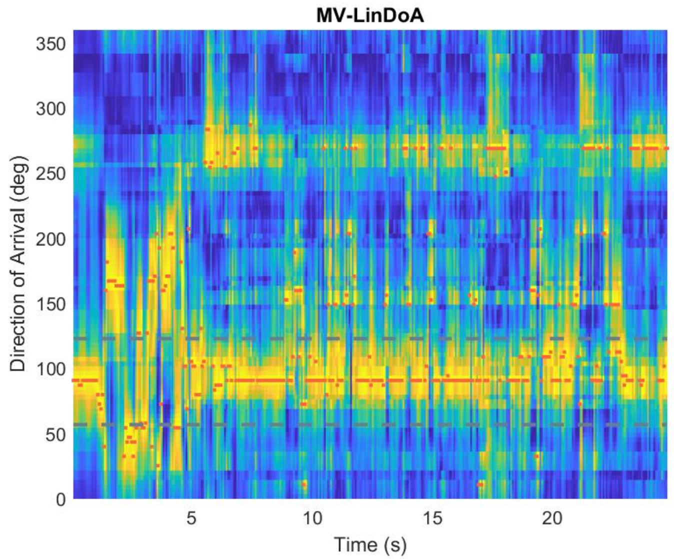


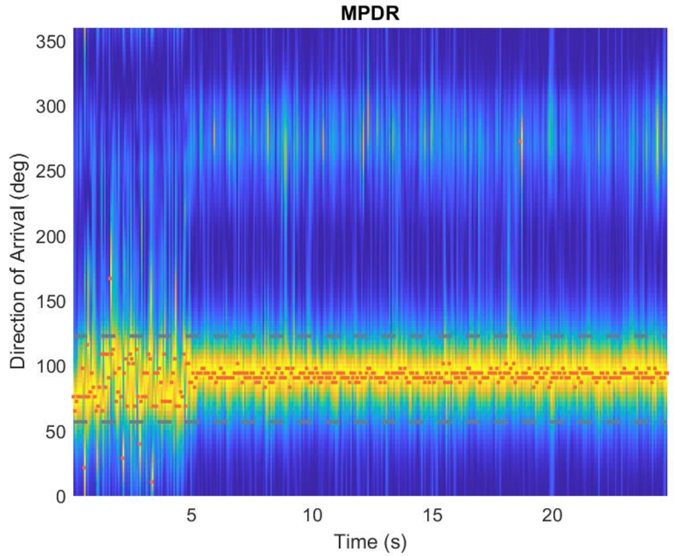


Hardware Still Music

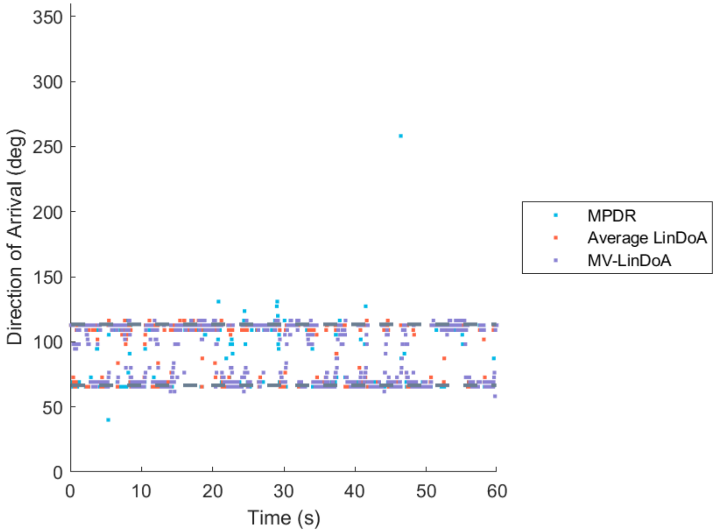


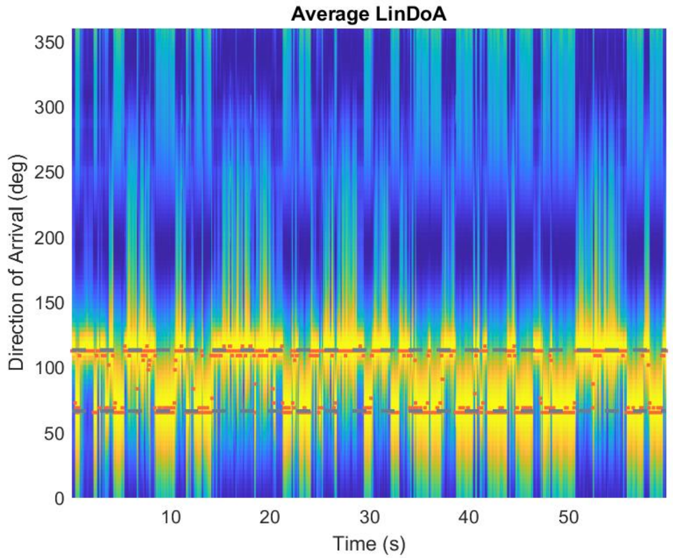


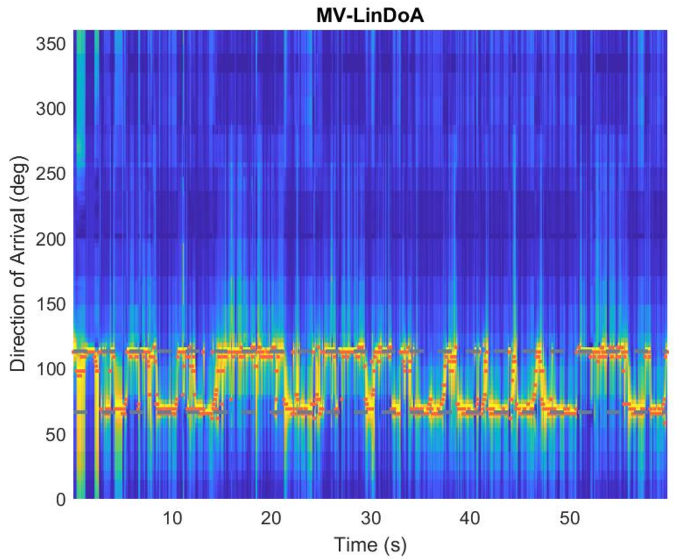


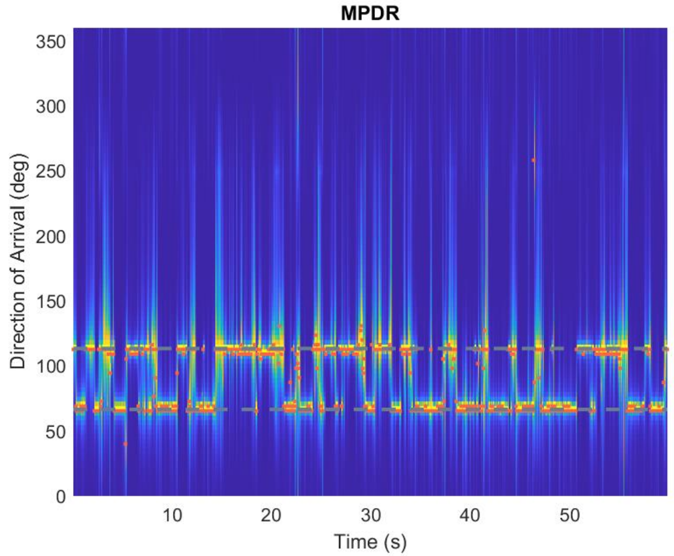


Simulation Still Dialogue

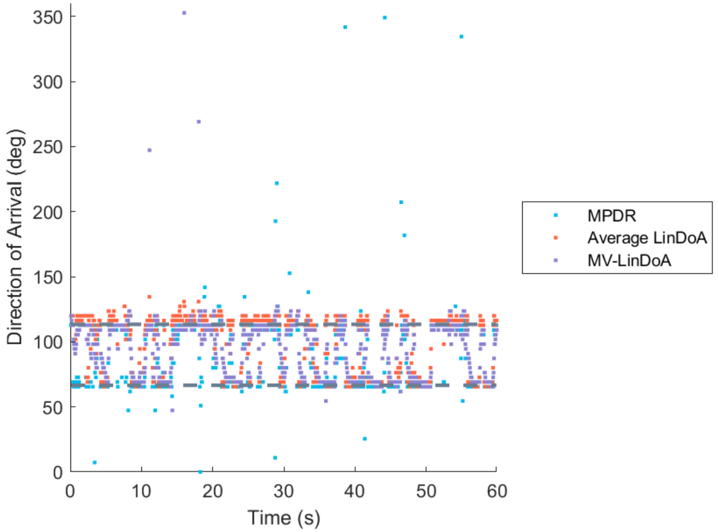


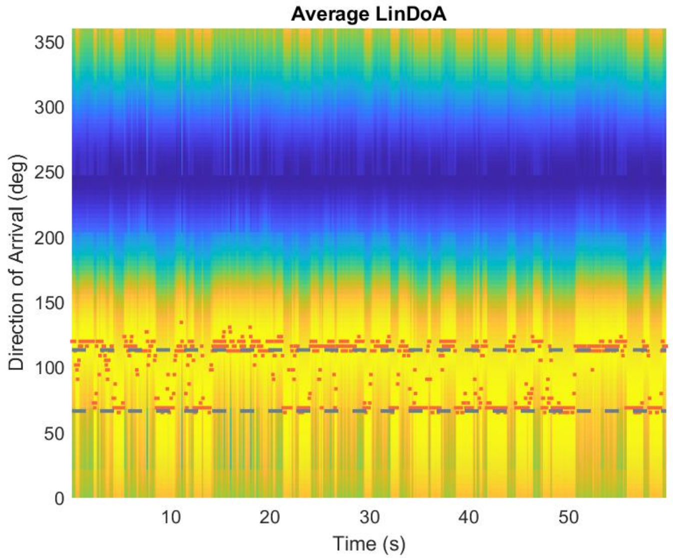


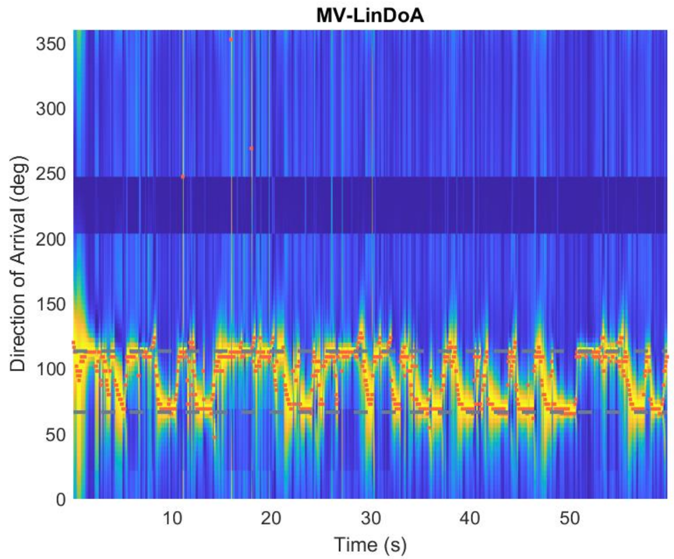


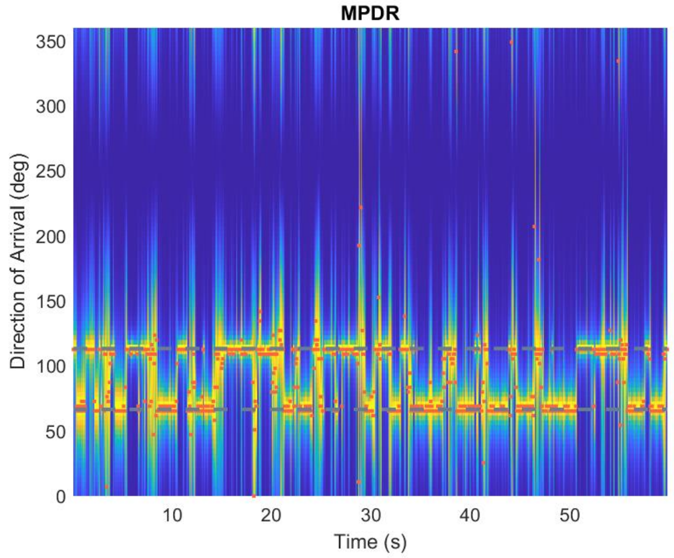


Simulation Still Small Dialogue

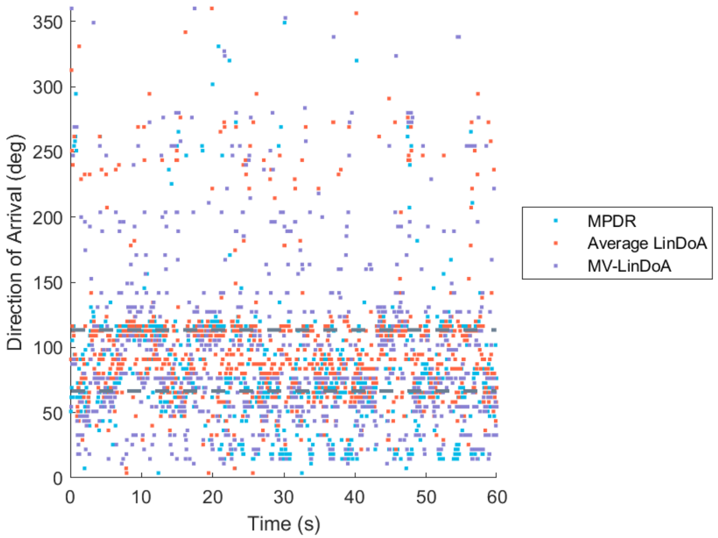


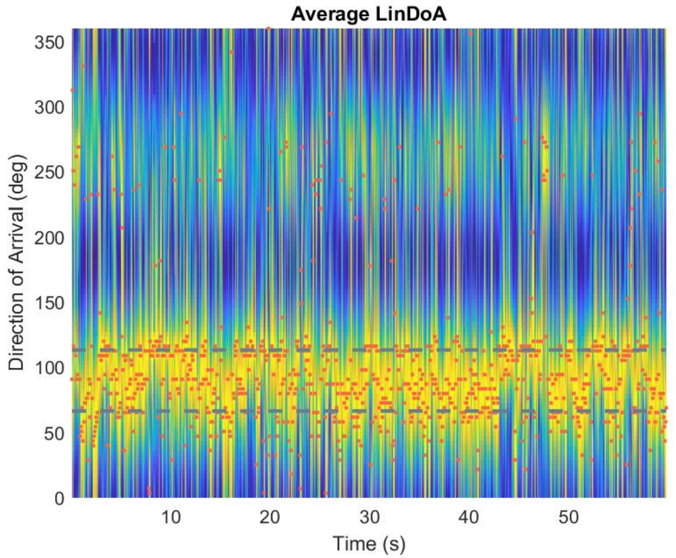


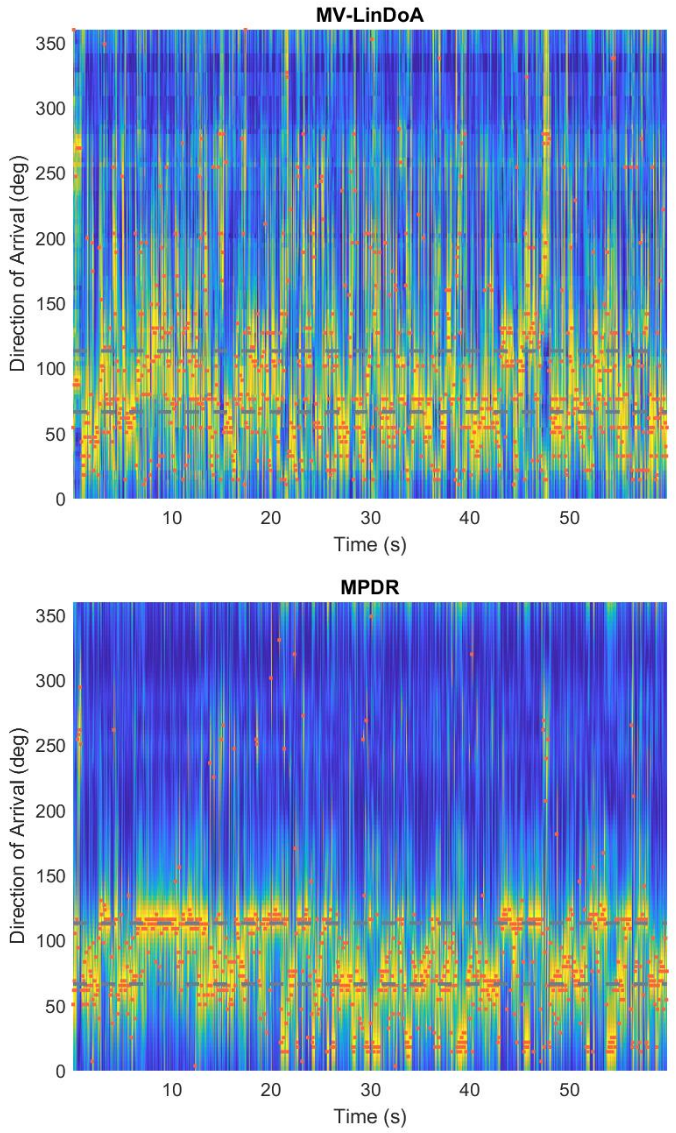




Hardware Still Dialogue







Beamforming

	Estimated Left	Estimated Right
True Left	0.8942	0.0975
True Right	0.1061	0.9534

Future Work

Compare to DS and MVDR

Array sizes and geometries

Parameters

Frequency

Beamforming

Power and Error

Summary

New method

Prototype

Experiments

Clas Veibäck

Thank you for listening!

LINKÖPING UNIVERSITY