Speech Intelligibility Enhancement using Microphone Array via Intra-Vehicular Beamforming

Final Presentation

Project By: Devin McDonald, Joseph Mesnard
Advised By: Dr. Yufeng Lu, Dr. In Soo Ahn

April 28, 2018

BRADLEY University

Agenda

- Problem Background
- Project Objectives
- Beamforming
- System Description
- Calibration
- Results
- Demo
- Future Work
- Questions

Problem Background

According to the National Safety Council, there are approximately

1.6 million

crashes **each year** due to distracted driving involving mobile phones [1].

Figure 1 - Man talking on phone while driving

BRADLEY University

Project Objectives

To reduce the risk of hands-on mobile phones usage in cars

- Increase speech intelligibility for far-end user
  - Uniform Linear Array (ULA) of microphones
  - Beamforming
  - Principle to Interference Signal Ratio

BRADLEY University

Problem Background

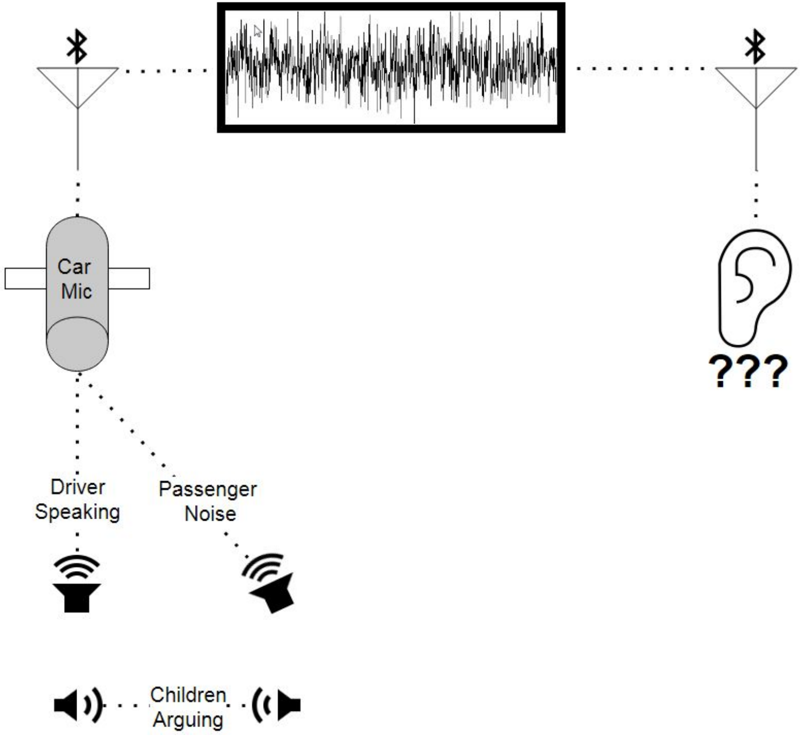


Figure 2 - Difficult to understand speech

Array of Microphones and Signal Processing

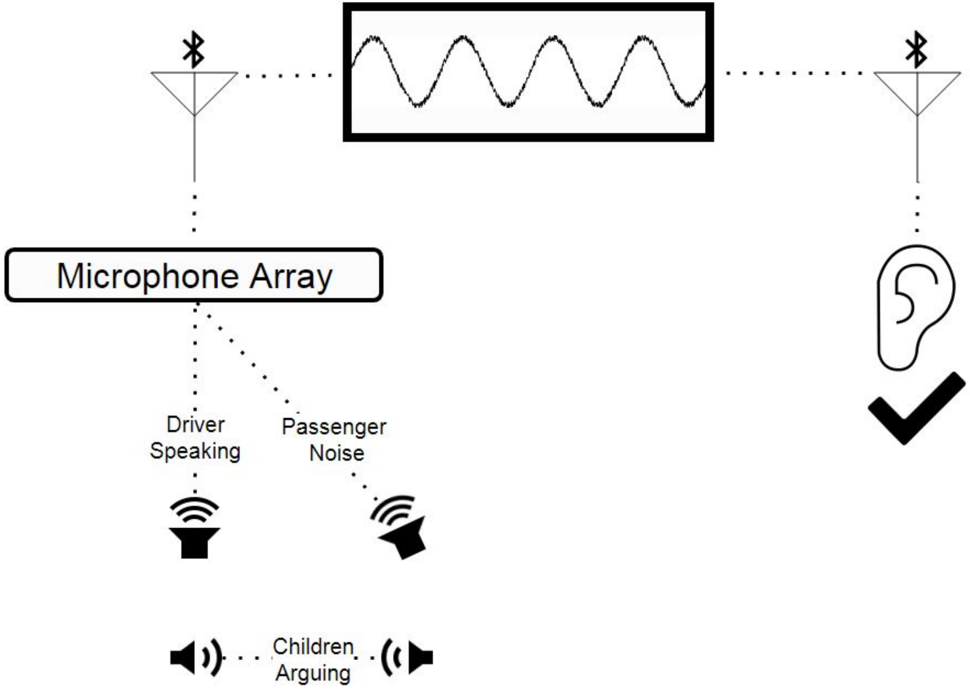


Figure 3 - Easier to understand speech

Microphone Array

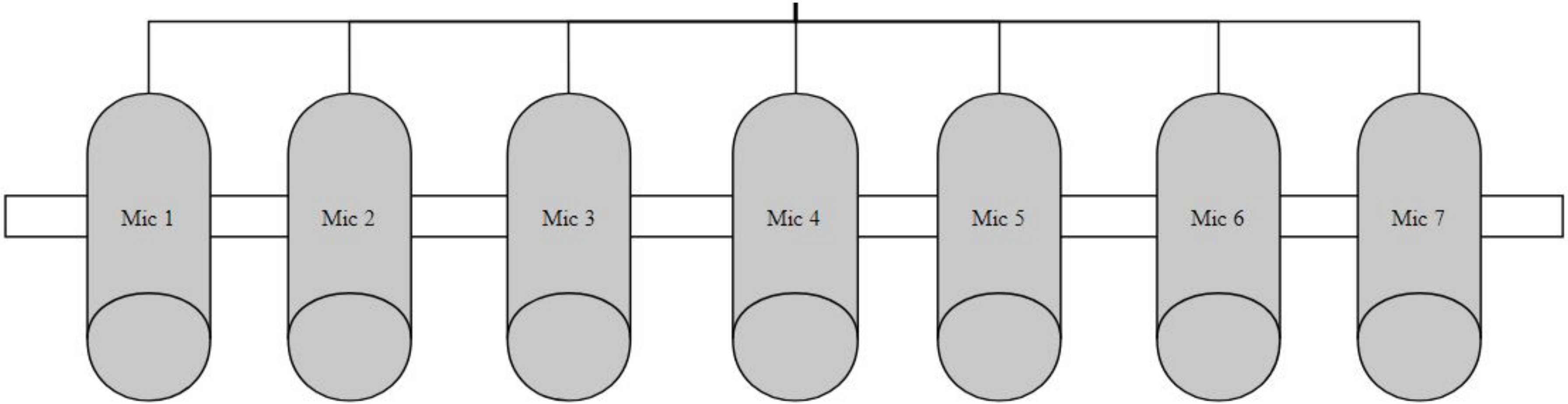


Figure 4 - Array design

Beamforming

- Beamforming or spatial filtering is a signal processing technique used in sensor arrays for directional signal transmission or reception.

- Delay-and-Sum Beamforming
  - Straightforward structure (see next few slides)
  - Simple implementation with less computation

Delay and Sum Beamforming

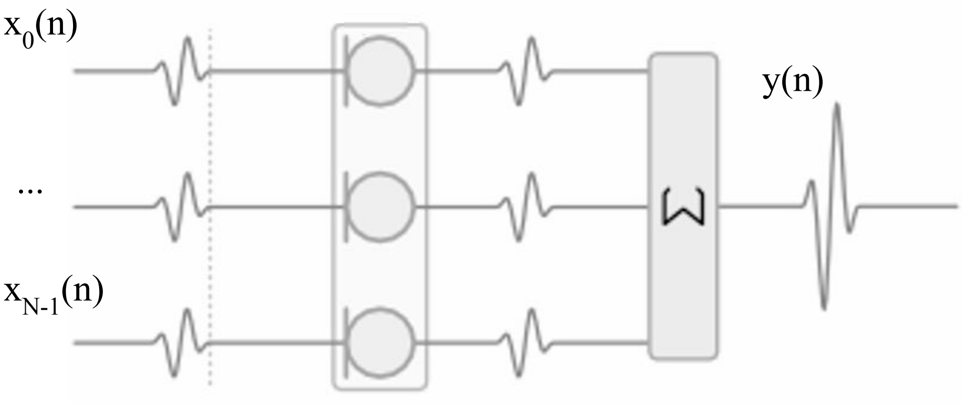


Figure 5 - Delay and Sum Beamforming at 0° explained [5]

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} x_k[n]$$

Delay and Sum Beamforming

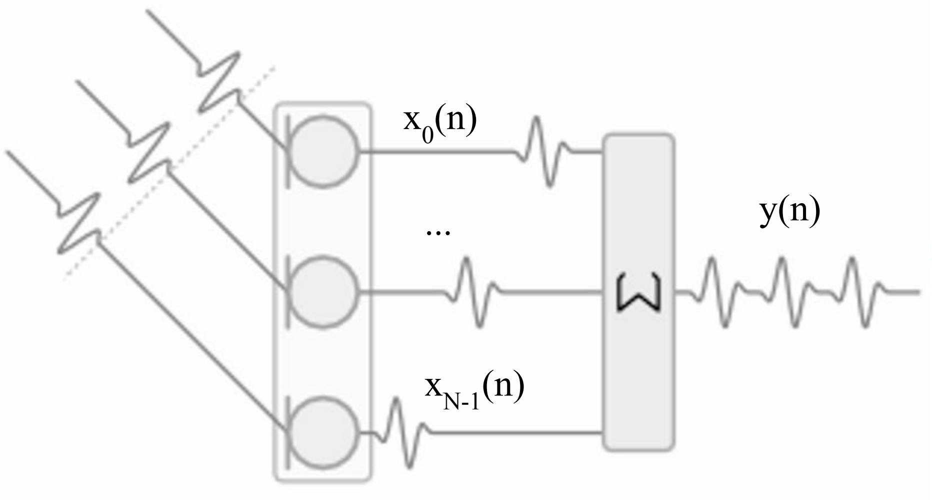


Figure 6 - Delay and Sum Beamforming at 45° explained [5]

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} x_k[n]$$

BRADLEY University

Delay and Sum Beamforming

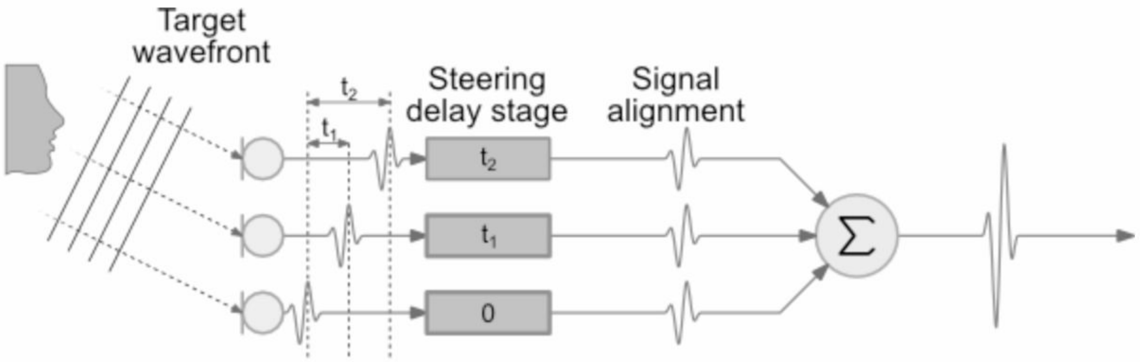


Figure 7 - Delay and Sum Beamforming with delays [5]

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} x_k[n - k\tau]$$

$\tau \rightarrow$ unit time delay

Requirements

Functional

- ❏ The system includes a ULA microphone array.
- ❏ Each microphone is routed to a system (such as MATLAB) for data acquisition.
- ❏ Beamforming is implemented in real-time.

Non-Functional

- ❏ The system will increase the intelligibility of near-end speech sent to the far-end user.
- ❏ The system requires little user manipulation or calibration.
- ❏ The system can be integrated within a vehicle.

System Block Diagram

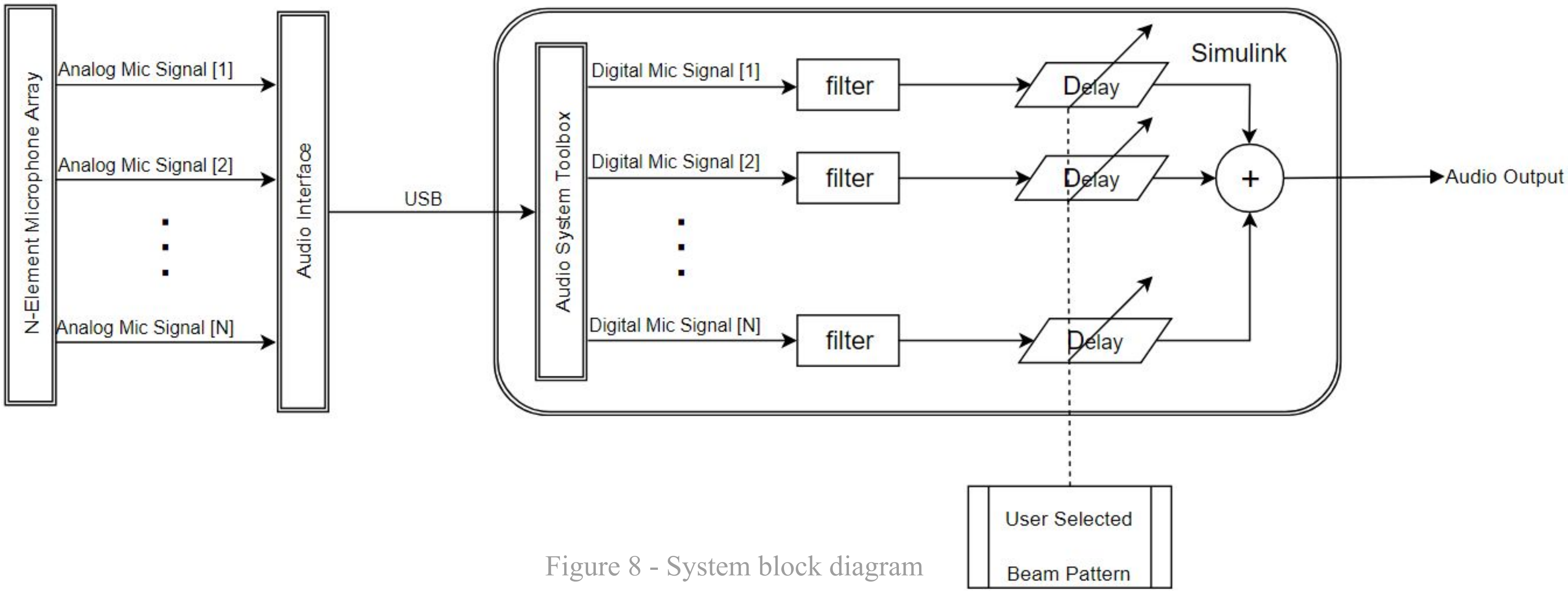


Figure 8 - System block diagram

Software and Hardware

- Simulink
  - Mathworks application used to implement microphone input
- Audio System Toolbox
  - Toolbox inside of Simulink to input microphone data from interface
- Interface
  - Scarlett 18i20 digital microphone interface to attach microphones to
- Microphones
  - Cardioid polar pattern microphones
- Speaker
  - A speaker is used for calibration

Microphone Array Design

A linear microphone array is determined to be the best array design for this application

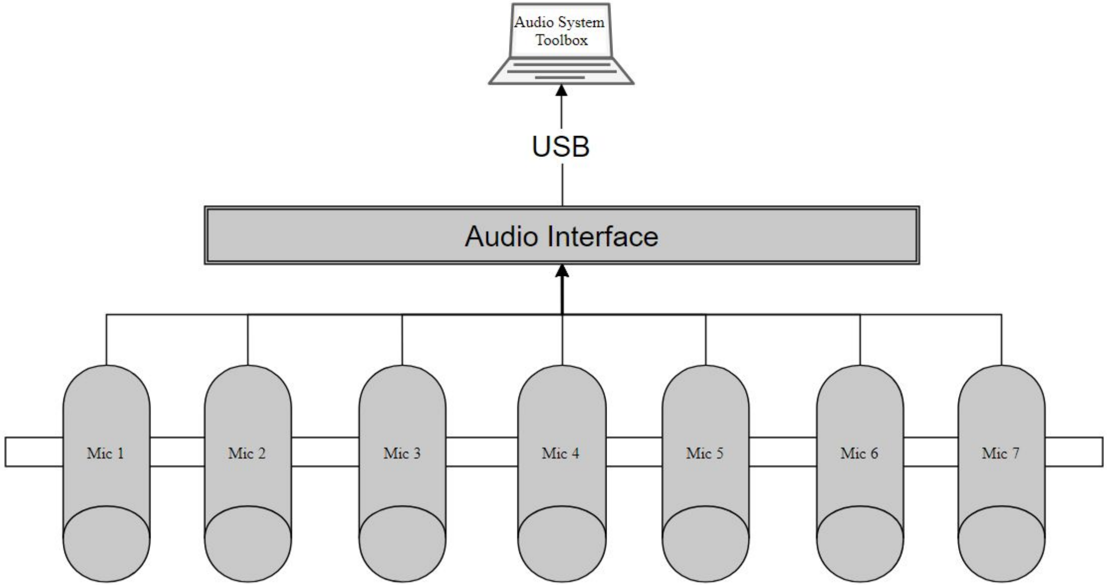


Figure 9 - Array design

Filtering

A-Weighting filters are used to focus on speech content

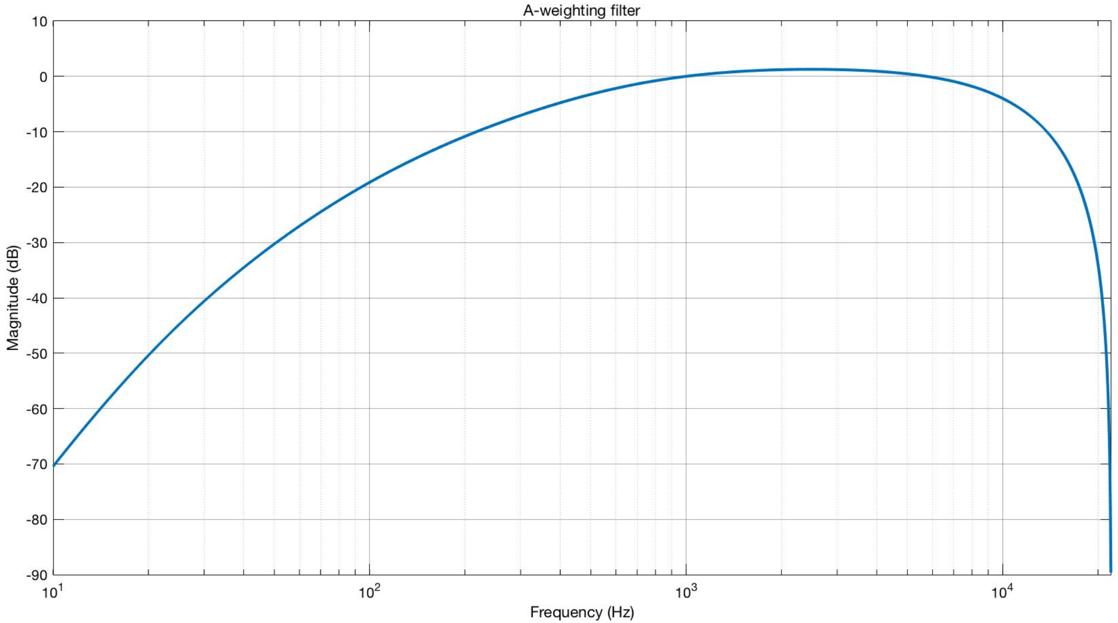


Figure 10 - A Weighting Filter

Fractional Delay

Fs = 44.1 kHz

f = 1 kHZ

Sampled sinc pulse

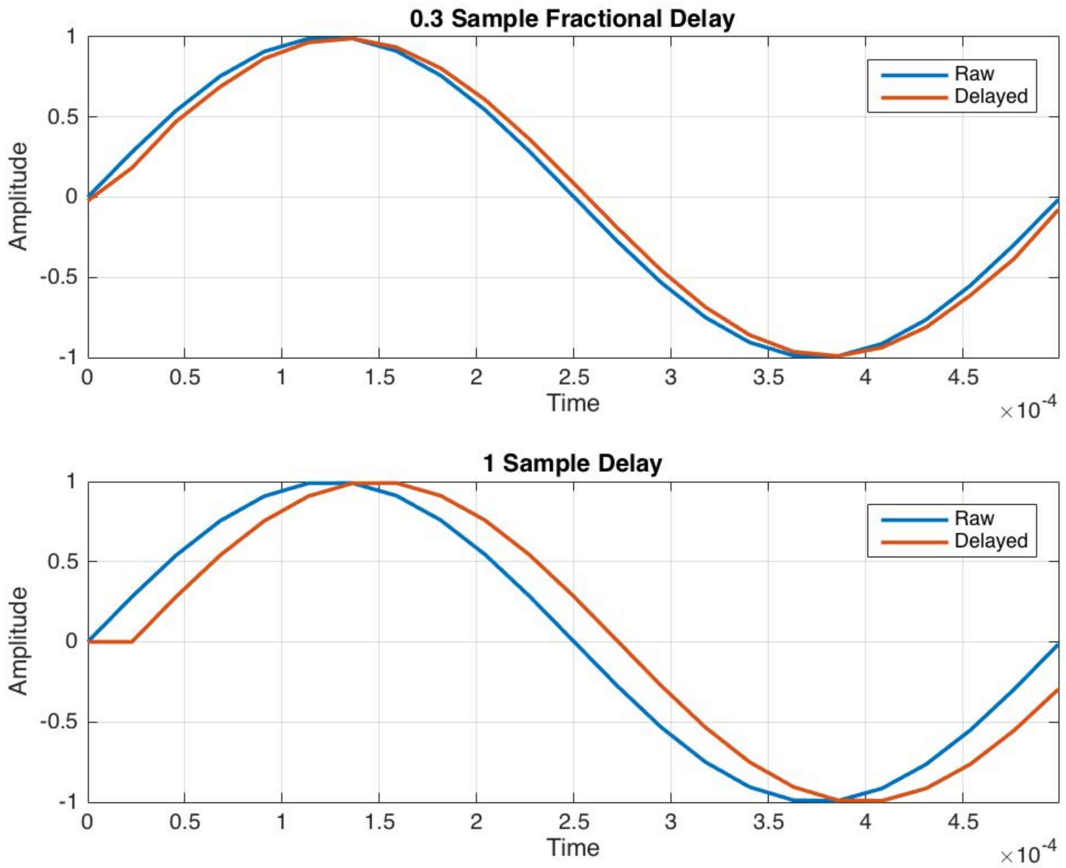


Figure 11 - Demonstration of fractional delays [5]

BRADLEY University

Fractional Delay

Achieved by sampling a sinc pulse to create a set of FIR filter coefficients

The sampling location is chosen based on the desired fractional delay

Higher number of sampled points creates a more accurate filter, but increases execution time

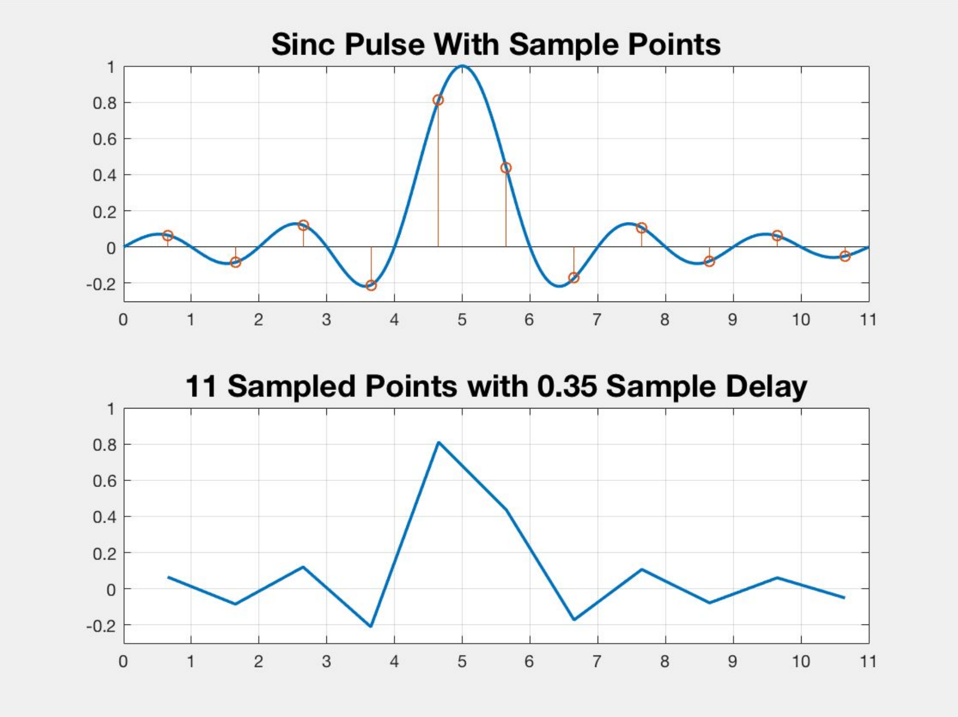


Figure 12 - Sinc pulse plot

BRADLEY University

Preliminary Results

Audio recorded using Logic Pro

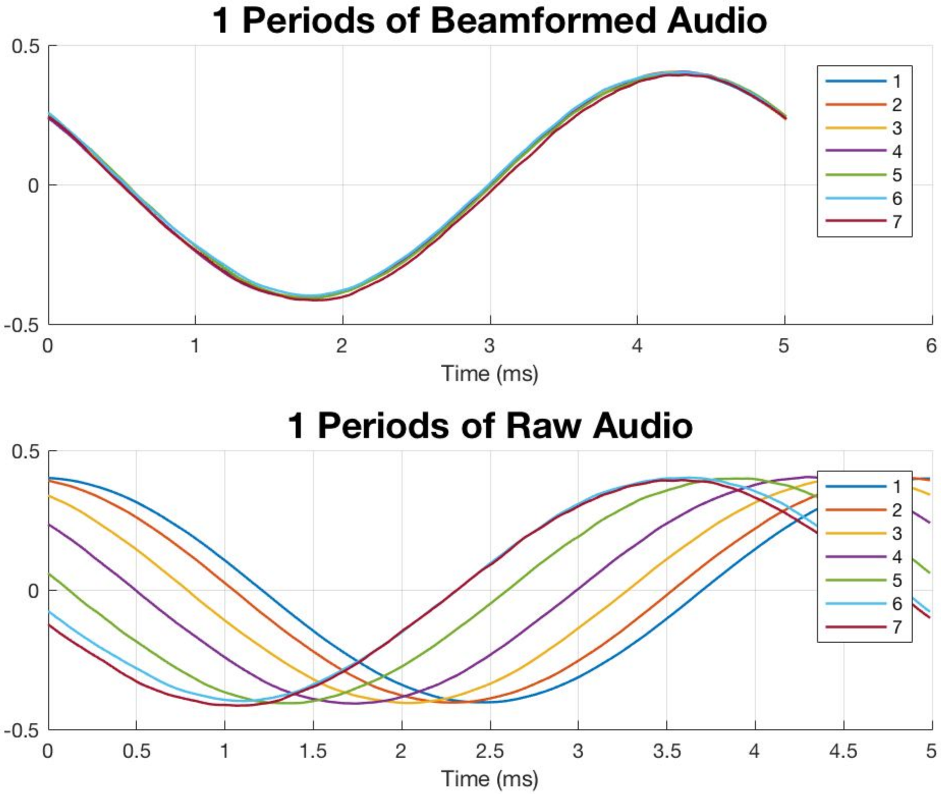


Figure 13 - Raw versus beamformed waves

Preliminary Results

```
Delay and sum beamforming took approximately: 2793.125ms for a 5.0s signal

RMS of sum of raw microphone audio: 1.567970
RMS of sum of beamformed audio:     1.988769
Total increase in signal RMS:       26.837226 %

Avg power of sum of raw microphone audio:        -25.975 dB
Avg power of sum of beamformed microphone audio: -21.842 dB
Total increase in dB:                             4.133 dB

Predominant Frequency (Beamformed) = 199.98 Hz

Predominant Frequency (Raw) = 200.04 Hz
```

Preliminary Results

Concerns

- Data sets recorded during the same tests in Logic contained different numbers of samples
- Initial tests used distance to calculate delay times

Calibration

Automatic Gain Controller is used to match the gain of the microphones

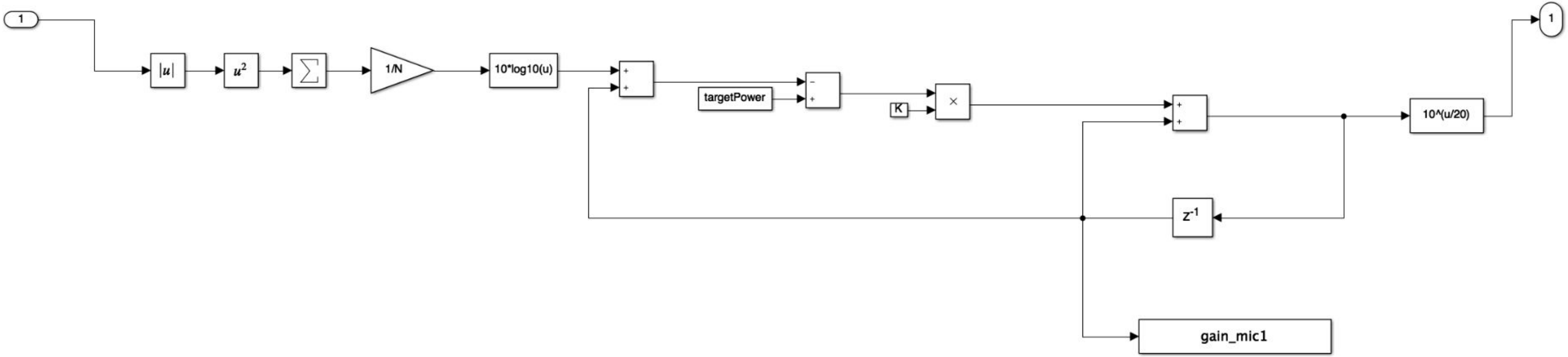


Figure 14 - AGC model for calibration

BRADLEY University

Calibration

The following Simulink model is used to calibrate the system

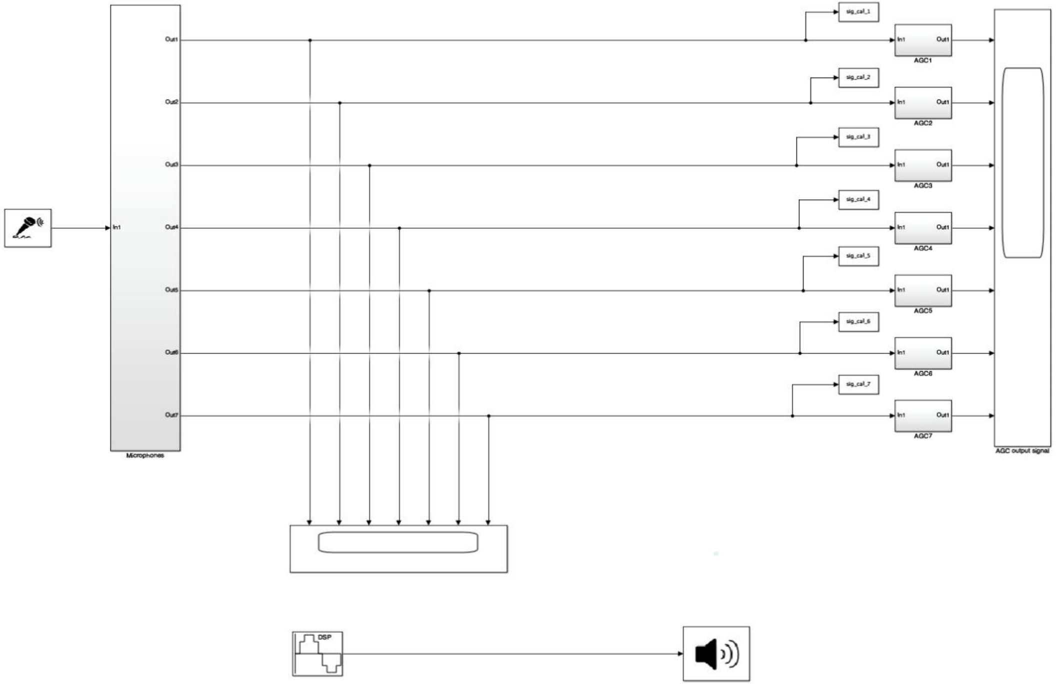


Figure 15 - Simulink calibration model

Calibration

A MATLAB Script calculates the time between zero crossings

Linear interpolation is used to calculate a precise zero crossing when it occurs between two samples

Plots are manually zoomed during calibration

Requires low frequency signal

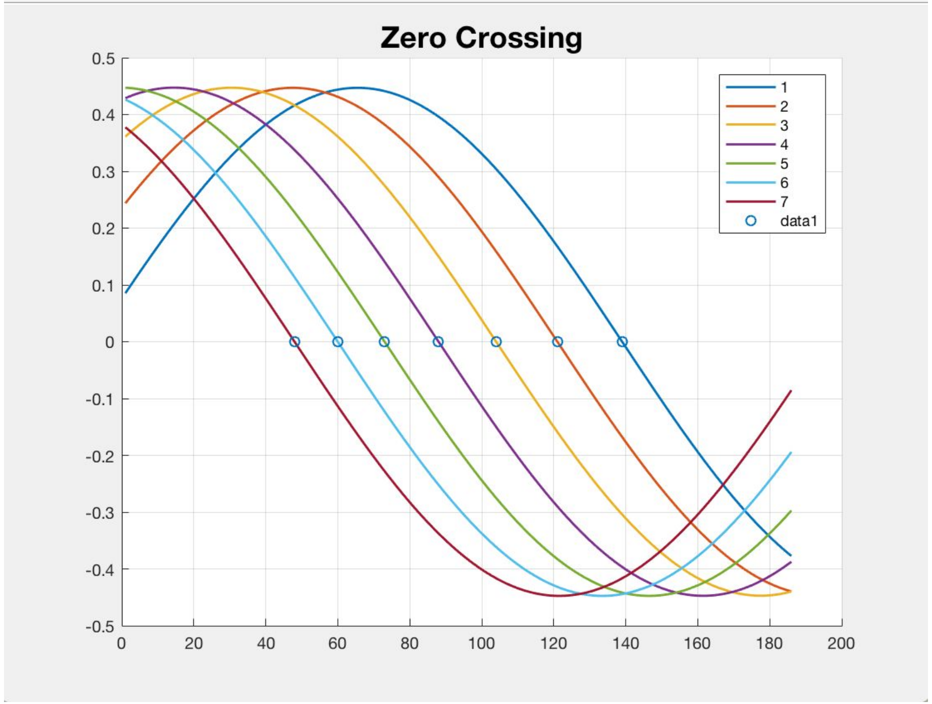


Figure 16 - Zero crossing of calibration signal

Calibration

The characteristics of the speaker system must be considered when calibrating the system.

- AGC
  - A 1 kHz sine wave must be played approximately at speaking level
- Delay Calculation
  - A speaker system with a good low frequency response is needed to calibrate the delays

Parts List

Quantity	Description	Price	Ext. Price
1	XLR Patch Cables	$31.75	$31.75
3	Behringer UltraVoice XM1800S Microphones	$39.99	$119.97
5	Pro Black Adjustable Dual Plastic 2pcs Drum Microphone Clip	$7.44	$37.20
1	Scarlett 18i20 Audio Interface	$499.99	$499.99

Simulation Calibration Input Subsystem

Uses a Simulink-generated sine wave instead of a microphone

Delay blocks are used to simulate physical delays

Gain blocks are used to simulate the different signal amplitudes caused by unmatched microphones and imprecise mixer gains

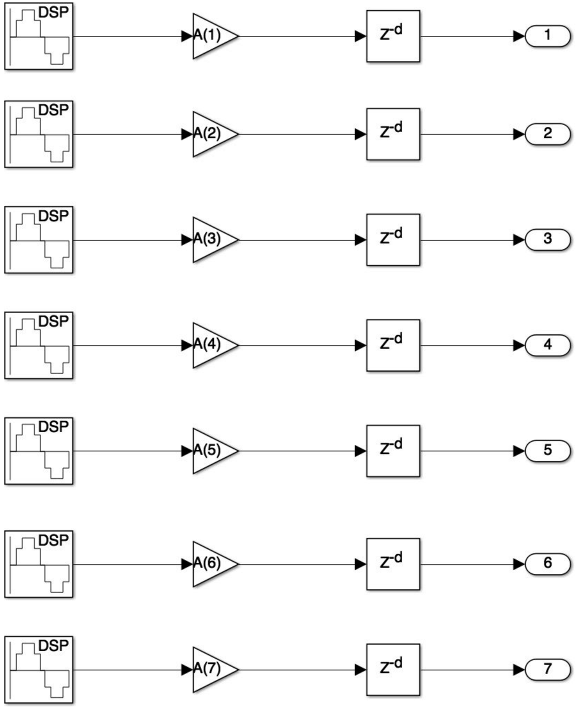


Figure 17 - Simulink calibration input

BRADLEY University

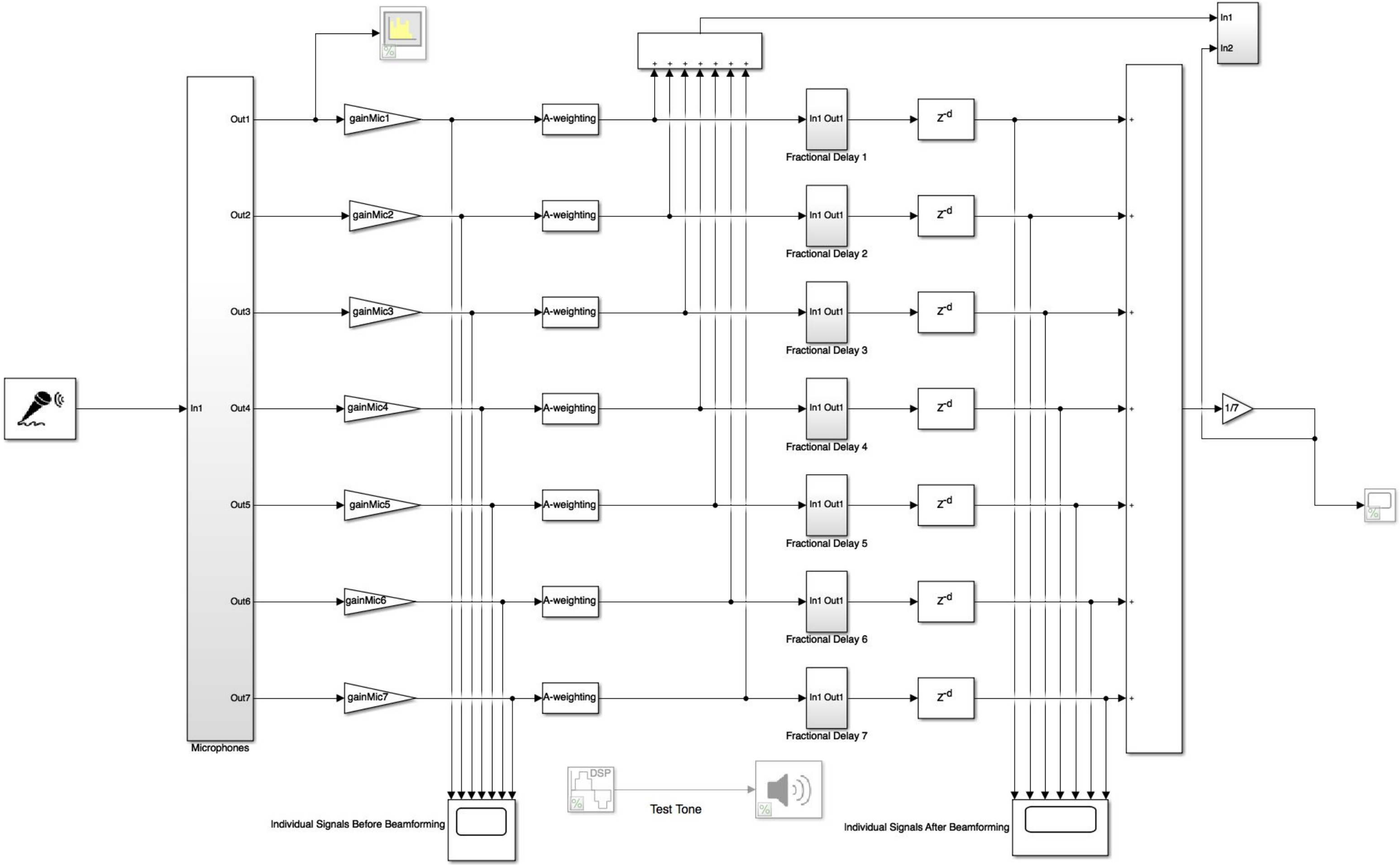


Figure 18 - Real-Time model

Simulation (400 Hz)

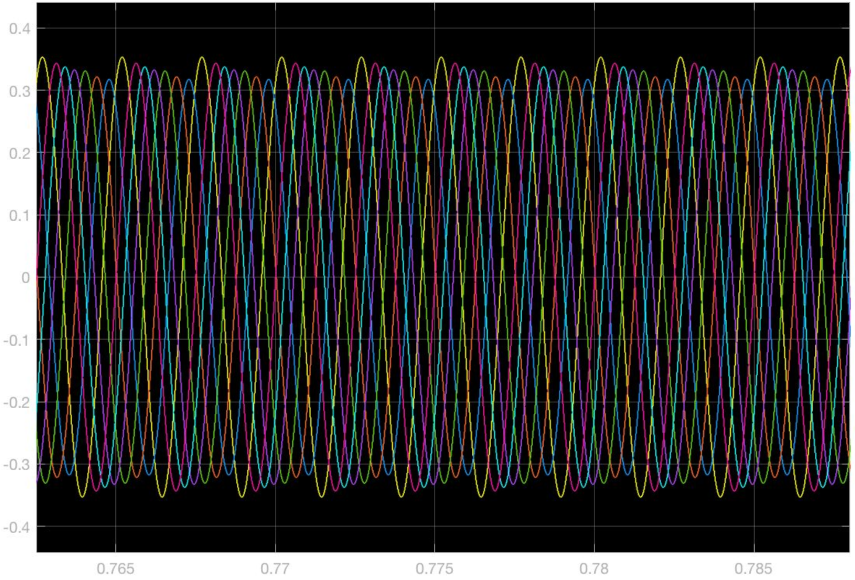


Figure 19. 400 Hz before beamforming

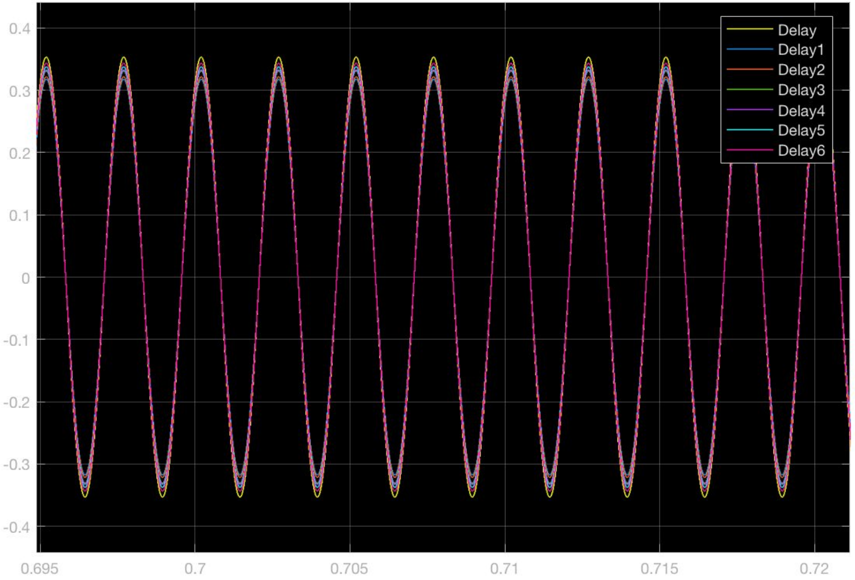


Figure 20. 400 Hz after beamforming

Simulation (400 Hz)

20.57 dB

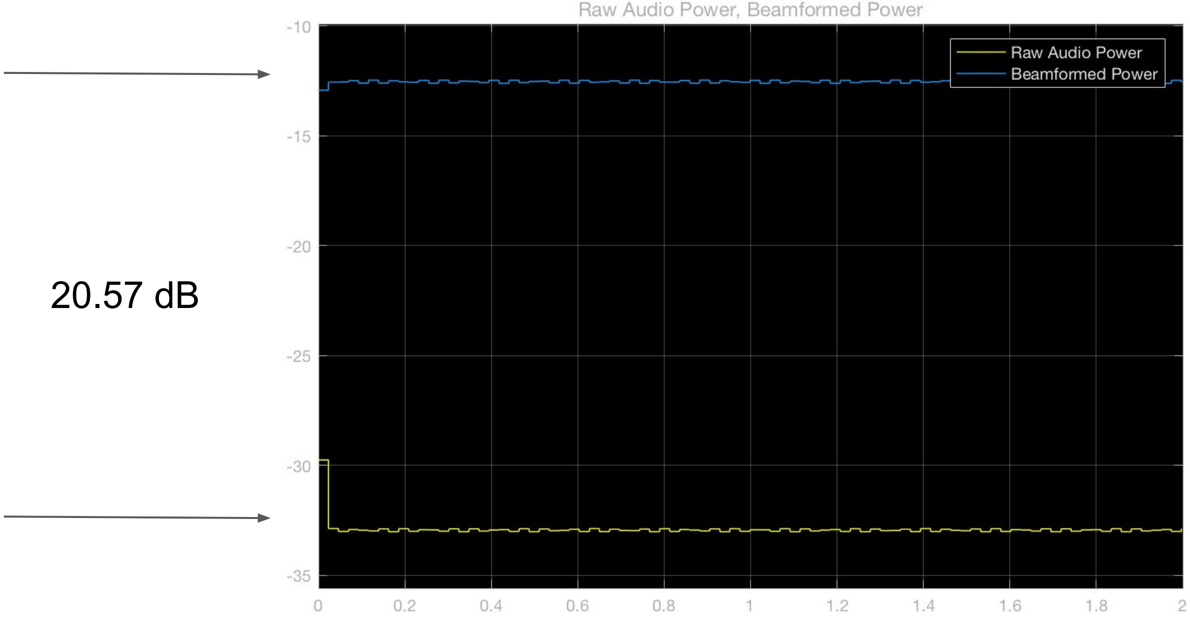


Figure 21. 400 Hz power plot

Simulation (1000 Hz)

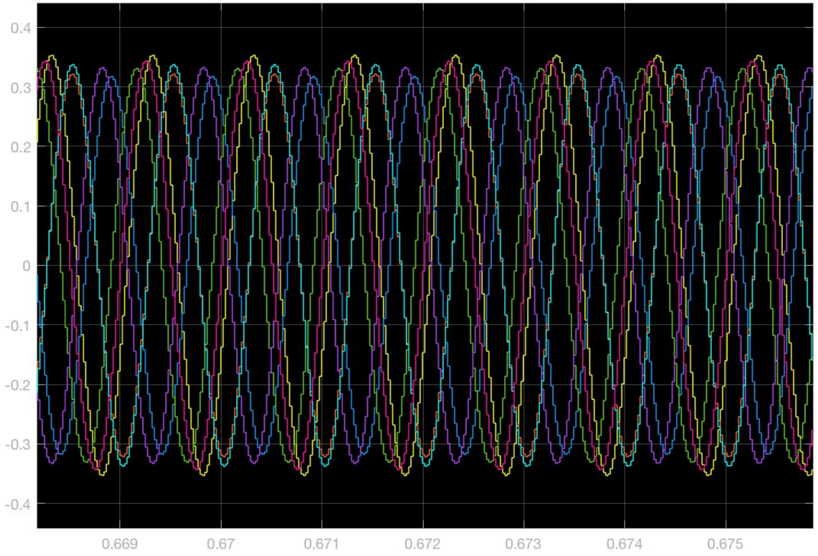


Figure 22. 1000 Hz before beamforming

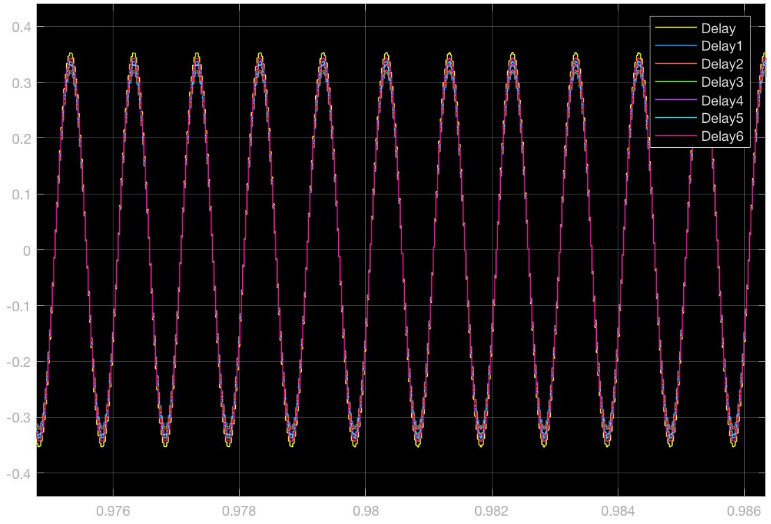


Figure 23. 1000 Hz after beamforming

BRADLEY University

Simulation (1000 Hz)

14.83 dB

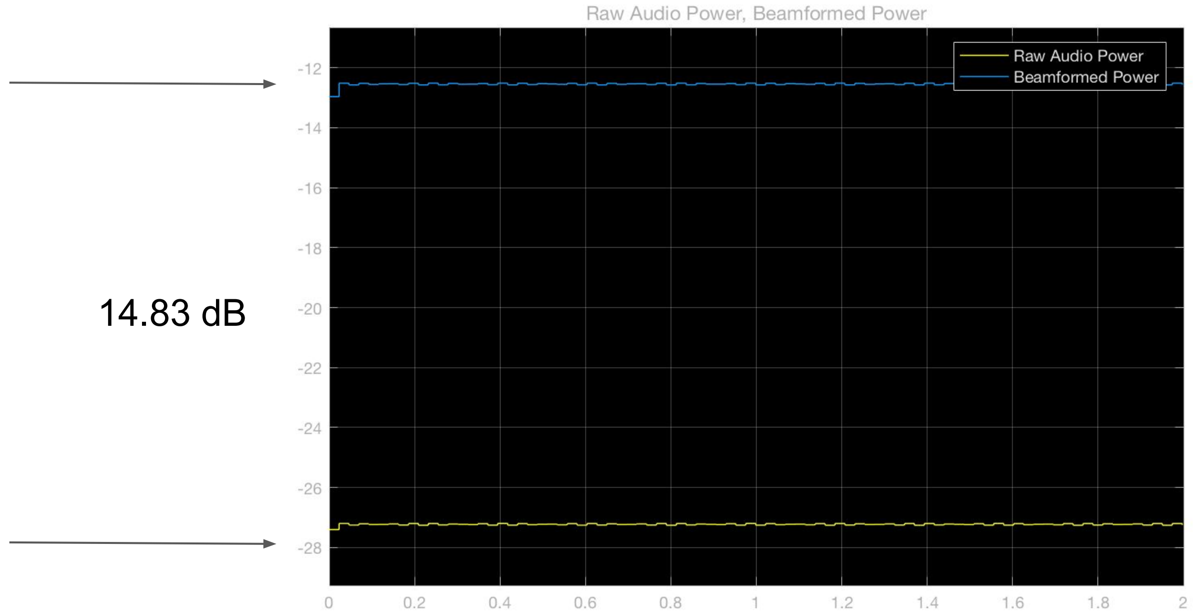


Figure X. 1000 Hz power plot

BRADLEY University

Simulation (3000 Hz)

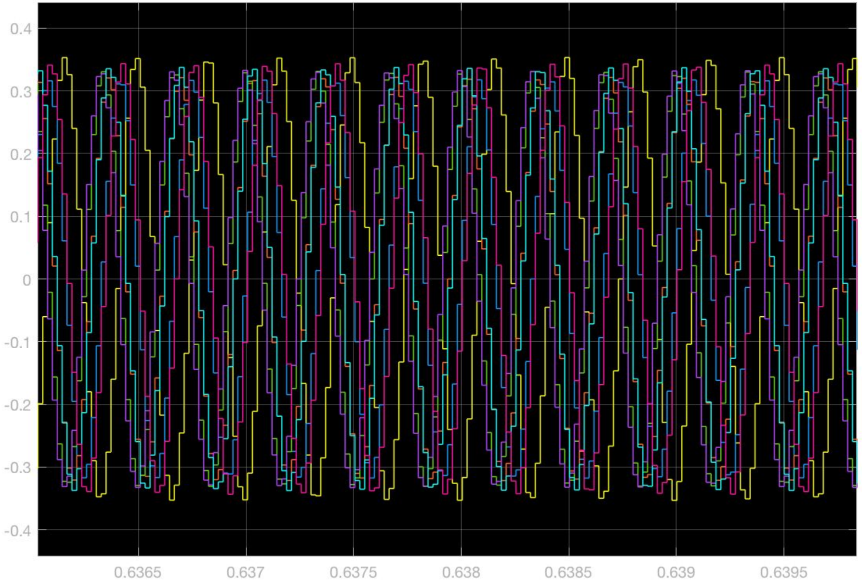


Figure 24. 3000 Hz before beamforming

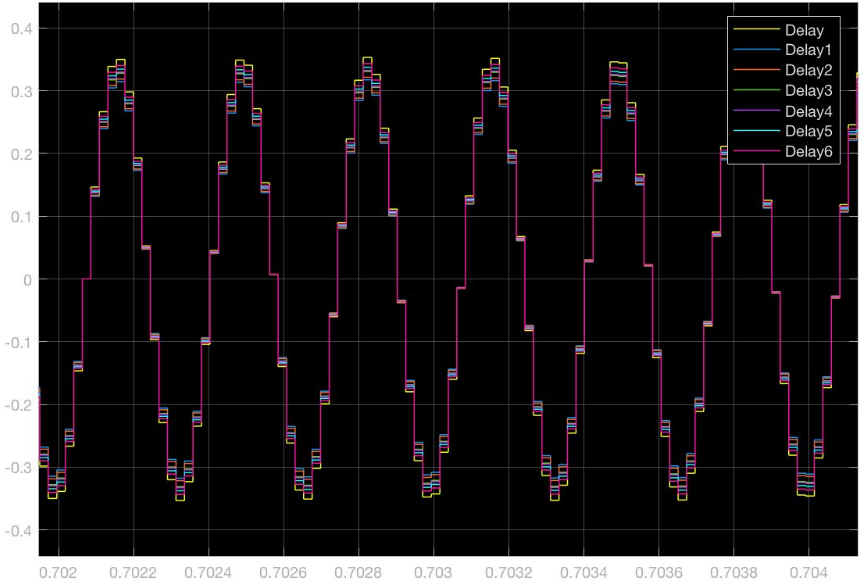


Figure 25. 3000 Hz after beamforming

BRADLEY University

Simulation (3000 Hz)

5.002 dB

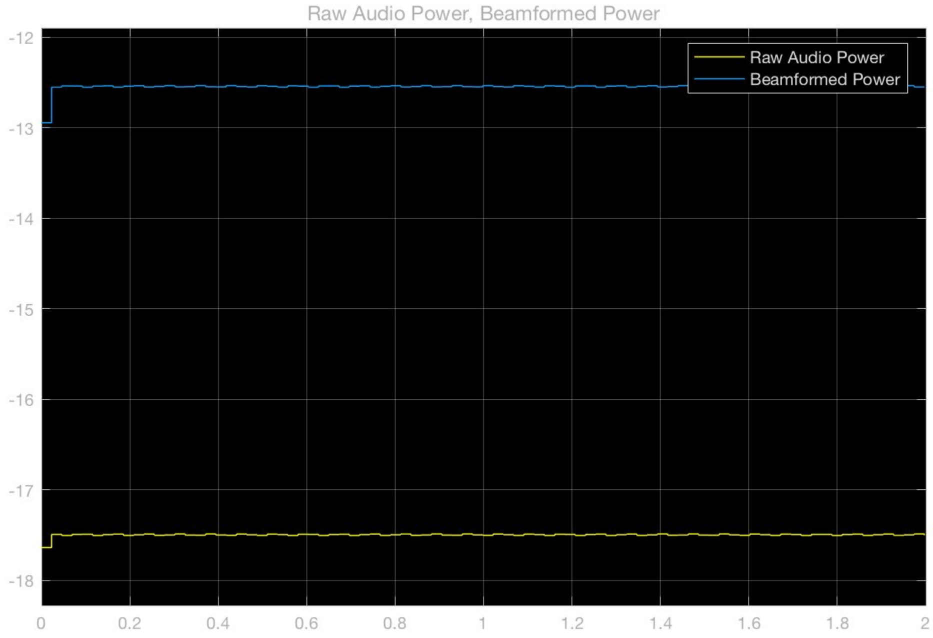


Figure 26. 3000 Hz power plot

Simulation (6000 Hz)

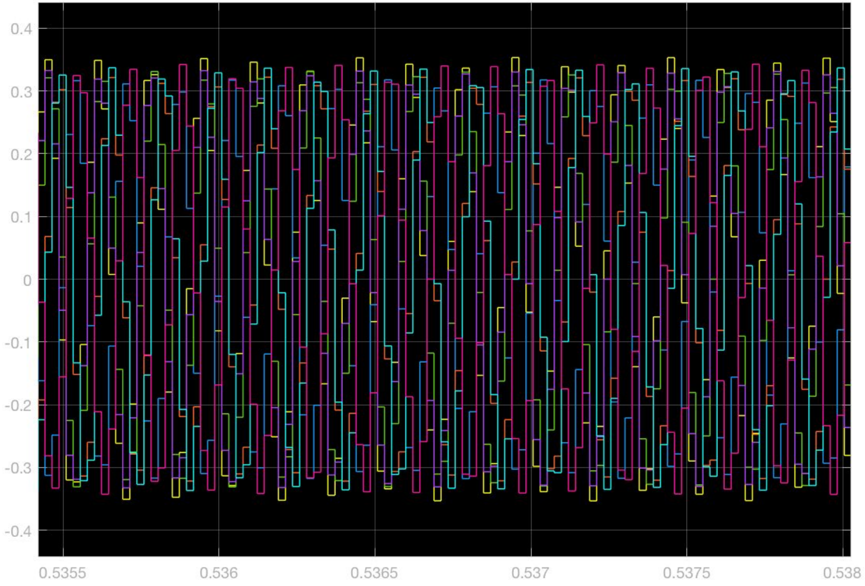


Figure 27. 6000 Hz before beamforming

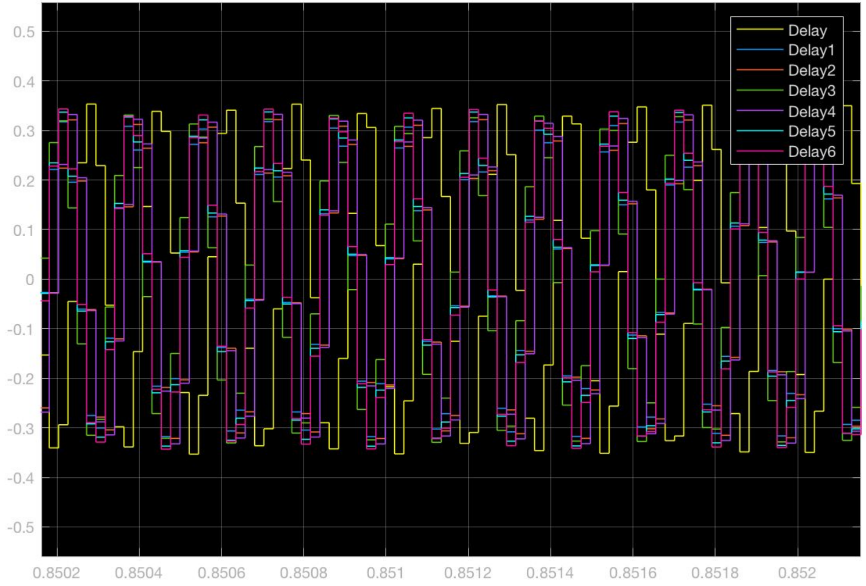


Figure 28. 6000 Hz after beamforming

BRADLEY University

Simulation (6000 Hz)

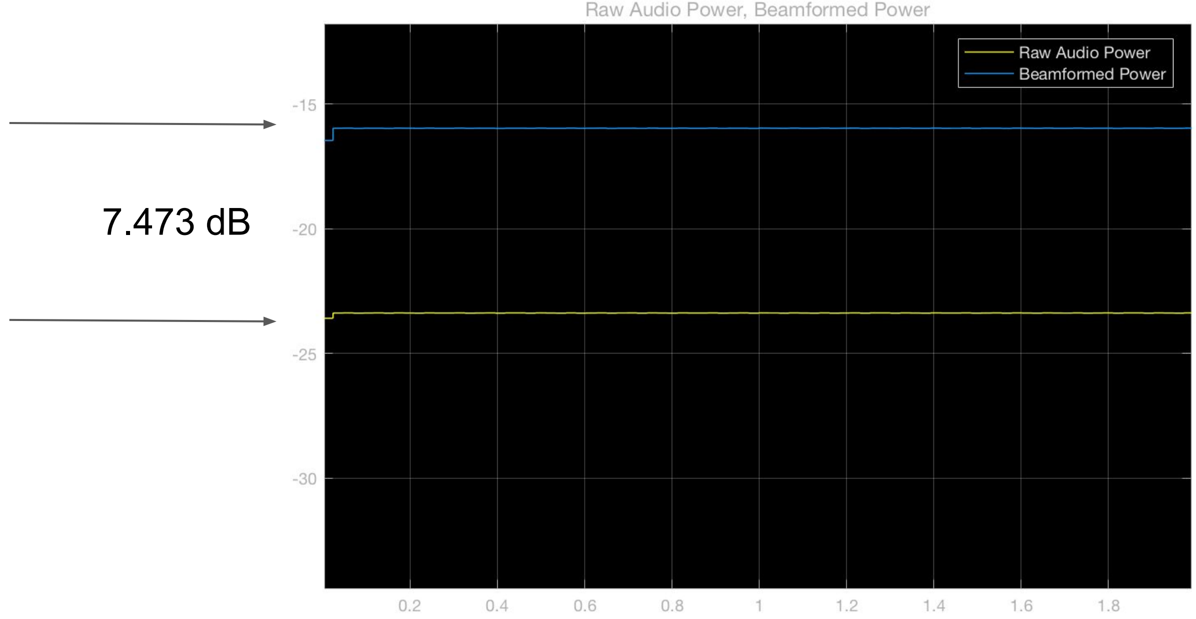
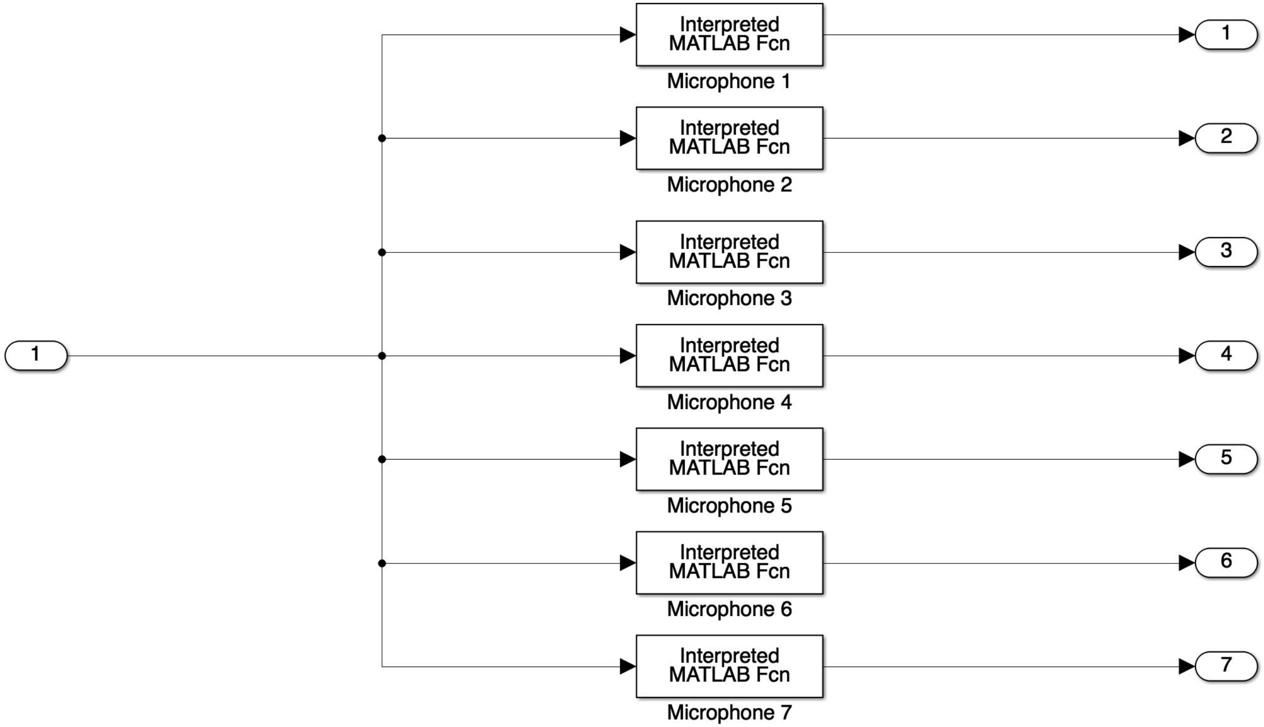


Figure 29. 6000 Hz power plot

BRADLEY University

Real-Time Input Subsystem



BRADLEY University

Figure 30 - Input system from interface

Results (400 Hz)

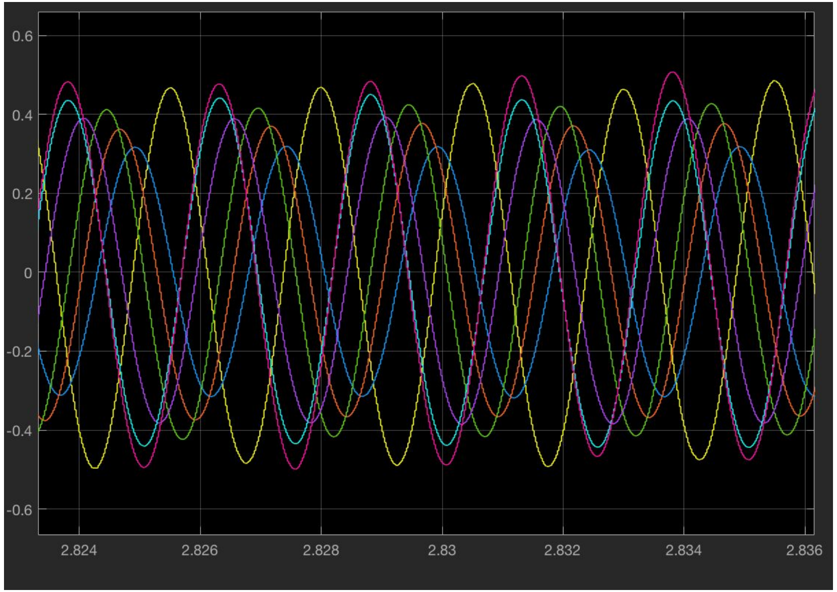


Figure 31. 400 Hz before beamforming

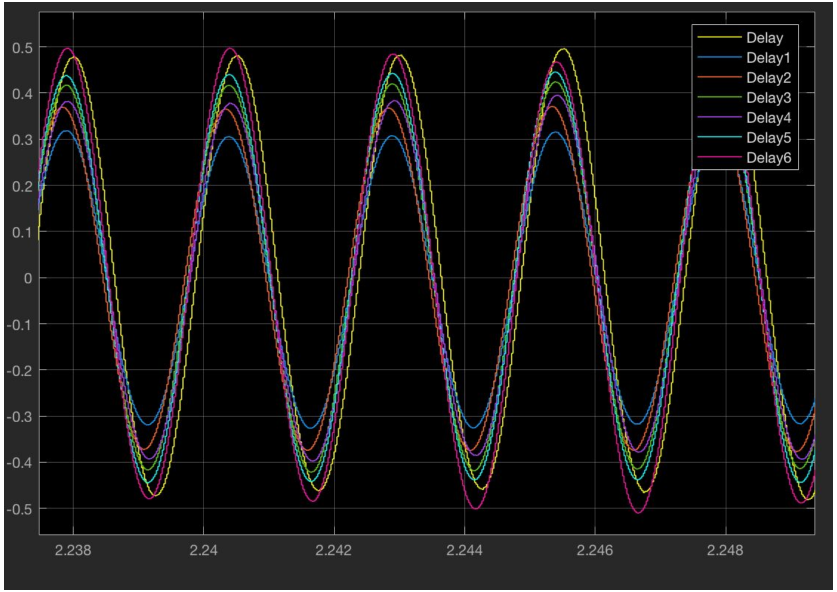


Figure 32. 400 Hz after beamforming

BRADLEY University

Results (400 Hz)

11.11 dB

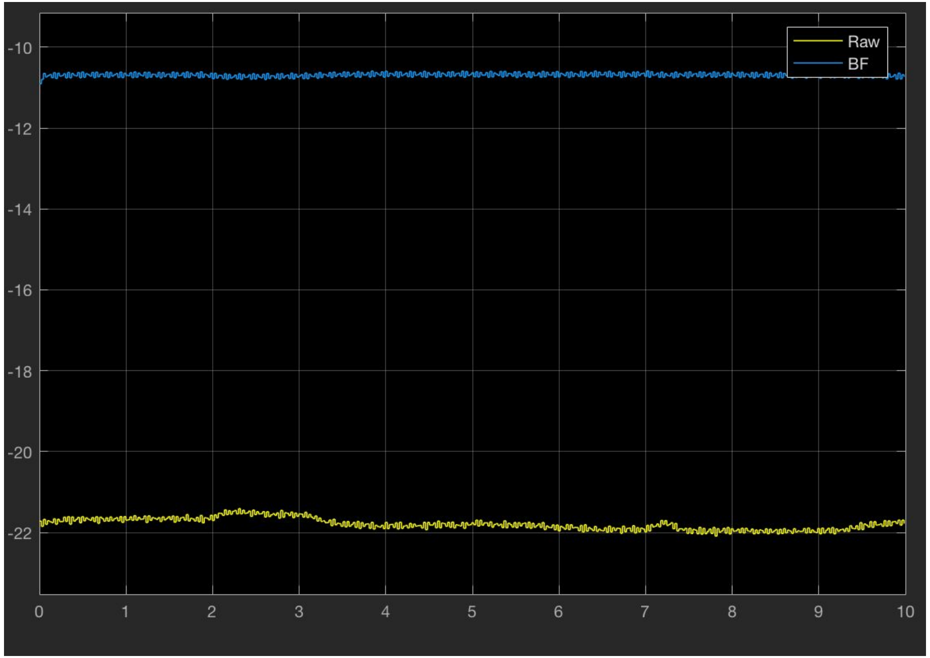


Figure 33. 400 Hz power plot

BRADLEY University

Results (1000 Hz)

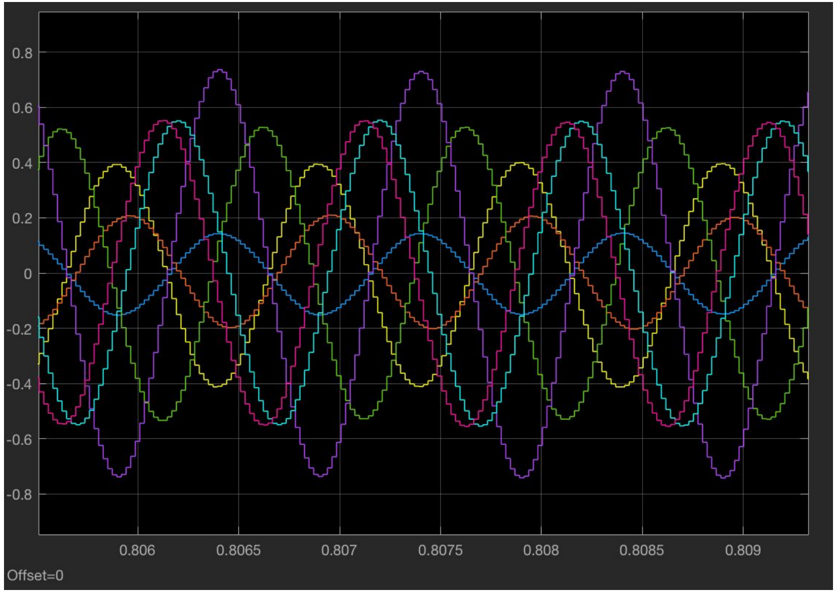


Figure 34. 1000 Hz before beamforming

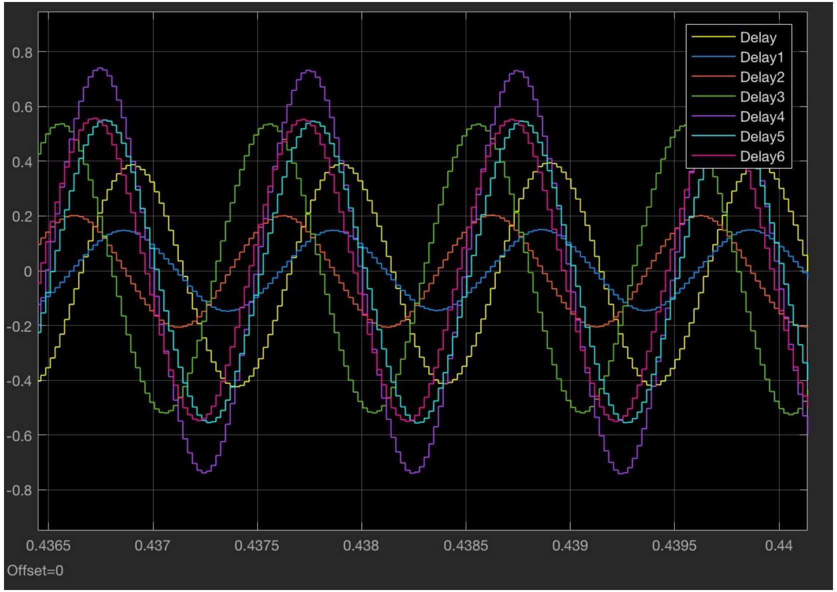


Figure 35. 1000 Hz after beamforming

BRADLEY University

Results (1000 Hz)

10.58 dB

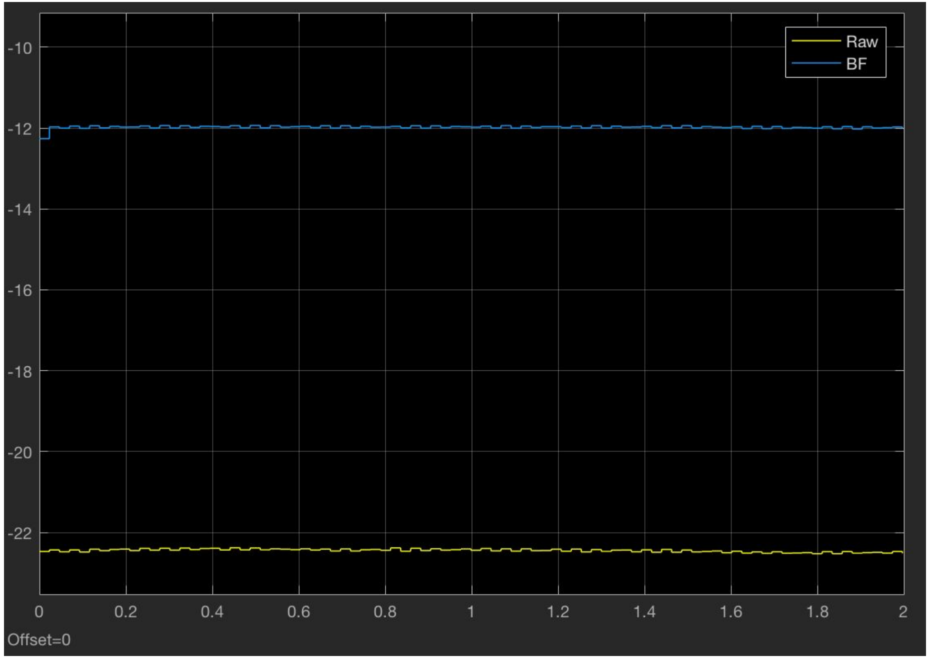


Figure 36. 1000 Hz power plot

BRADLEY University

Results (3000 Hz)

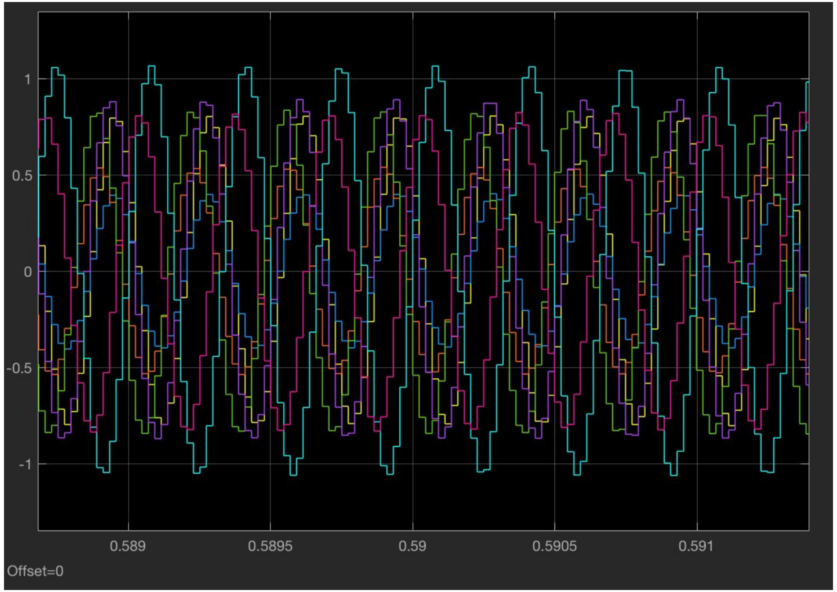


Figure 37. 3000 Hz before beamforming

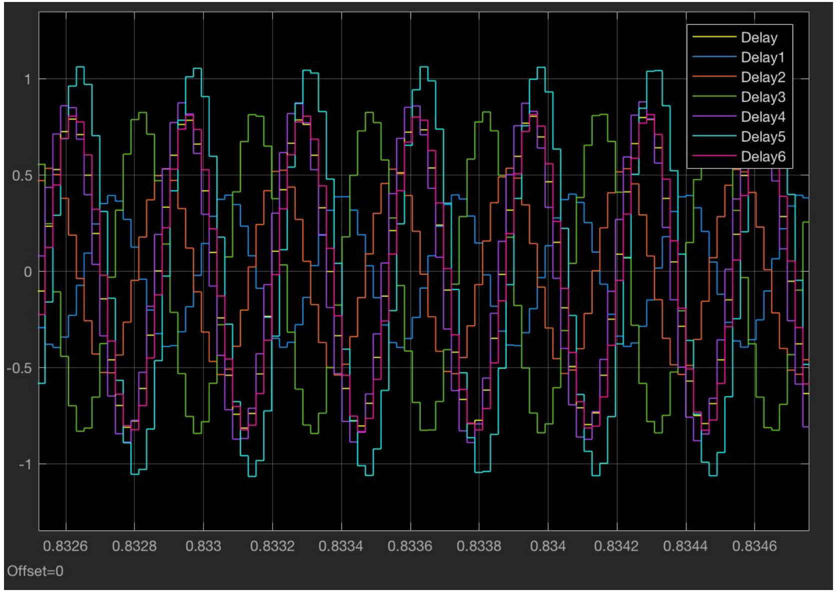


Figure 38. 3000 Hz after beamforming

BRADLEY University

Results (3000 Hz)

3.654 dB

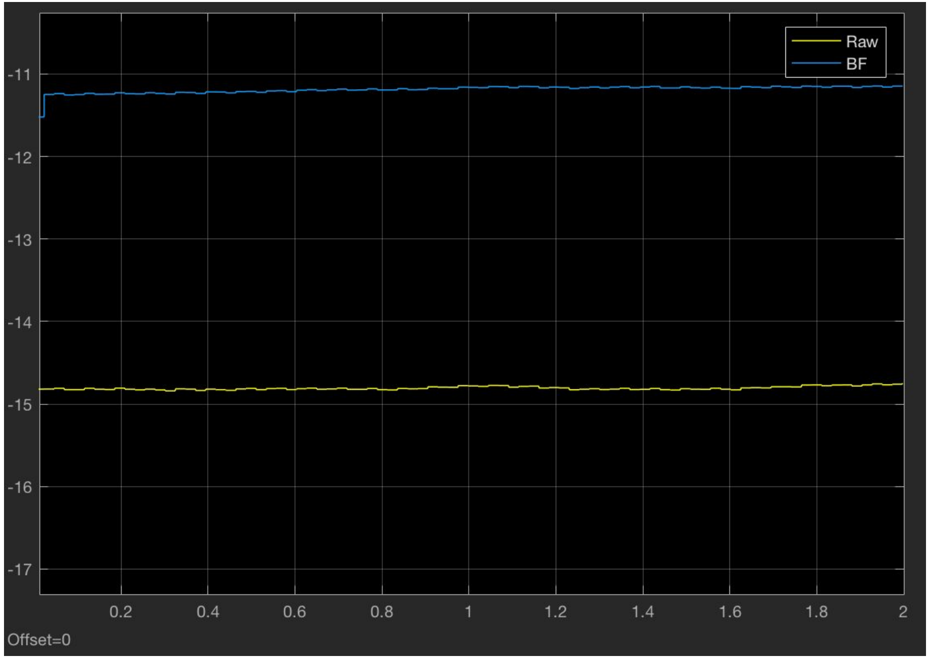


Figure 39. 3000 Hz power plot

BRADLEY University

Results (6000 Hz)

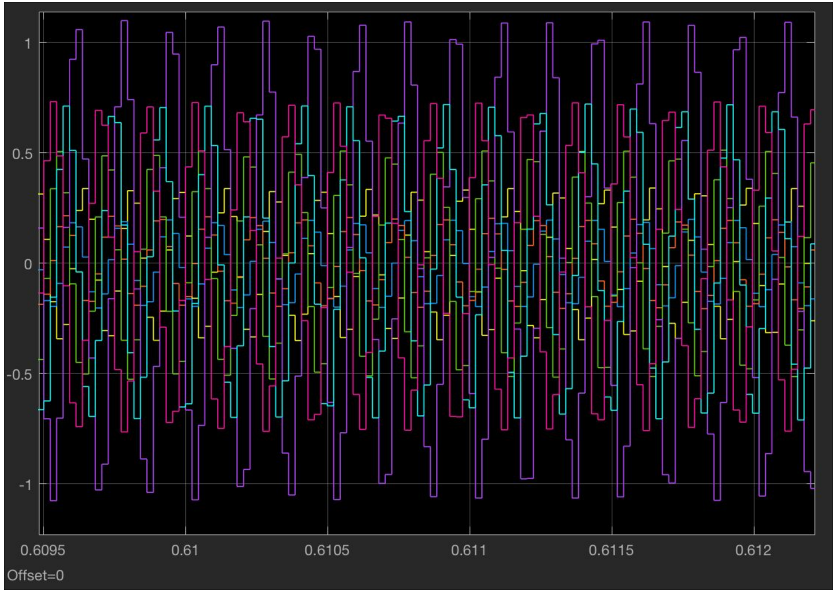


Figure 40. 6000 Hz before beamforming

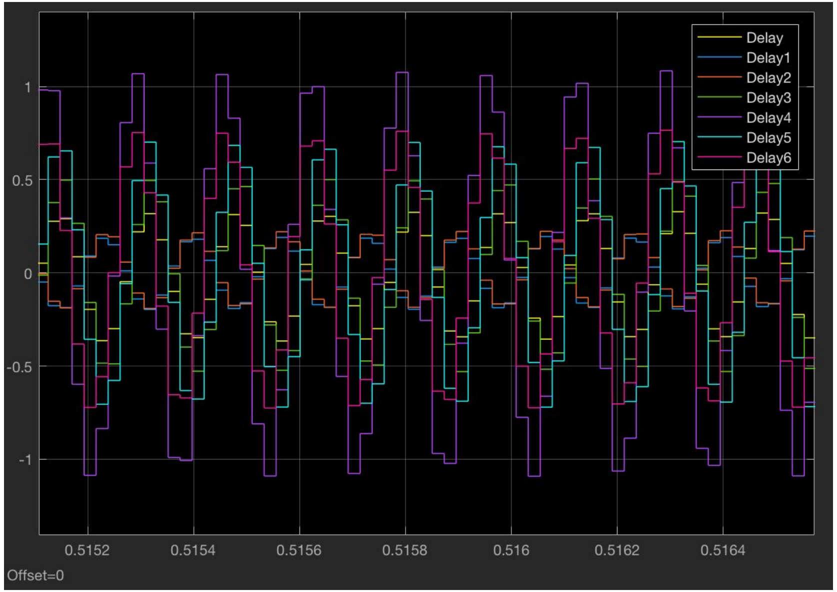


Figure 41. 6000 Hz after beamforming

Results (6000 Hz)

7.093 dB

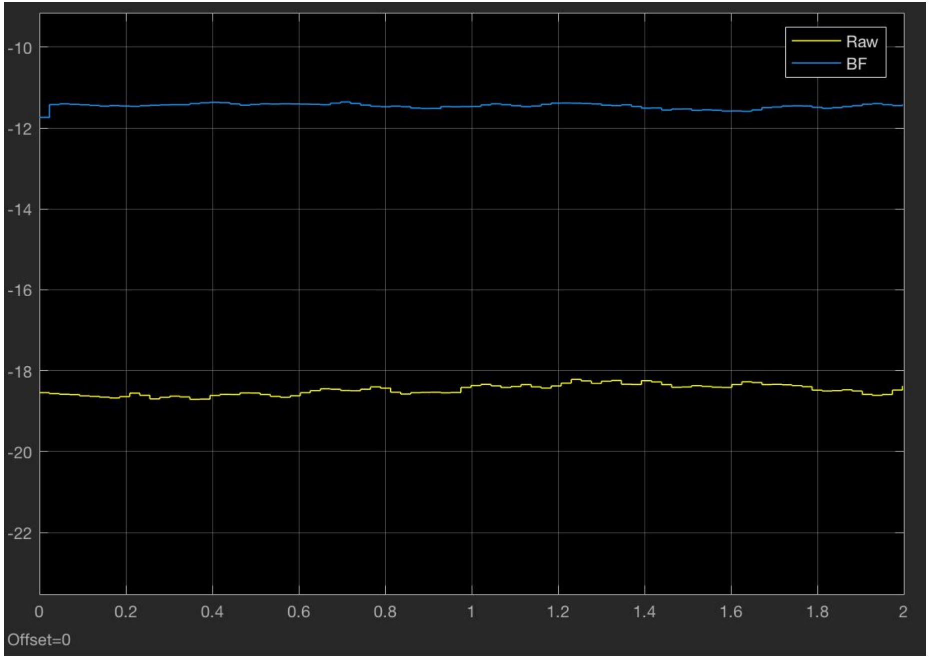
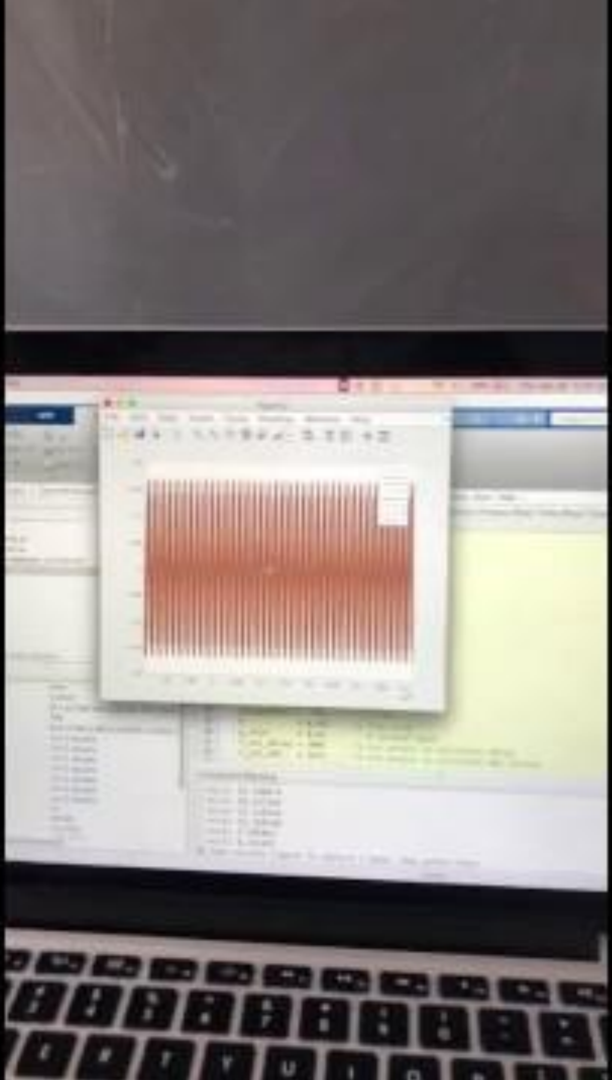


Figure 42. 6000 Hz power plot

Simulation Vs Real-Time Testing

Frequency	Simulation	Real-Time
400 Hz	20.57 dB	11.11 dB
1000 Hz	14.83 dB	10.58 dB
3000 Hz	5.002 dB	3.654 dB
6000 Hz	7.473 dB	7.093 dB

Calibration



Demo Audio

Before

After

Future Work

- Implement VAD into system
- Adaptive algorithm
- Non-linear array design

BRADLEY University

Engineering Efforts

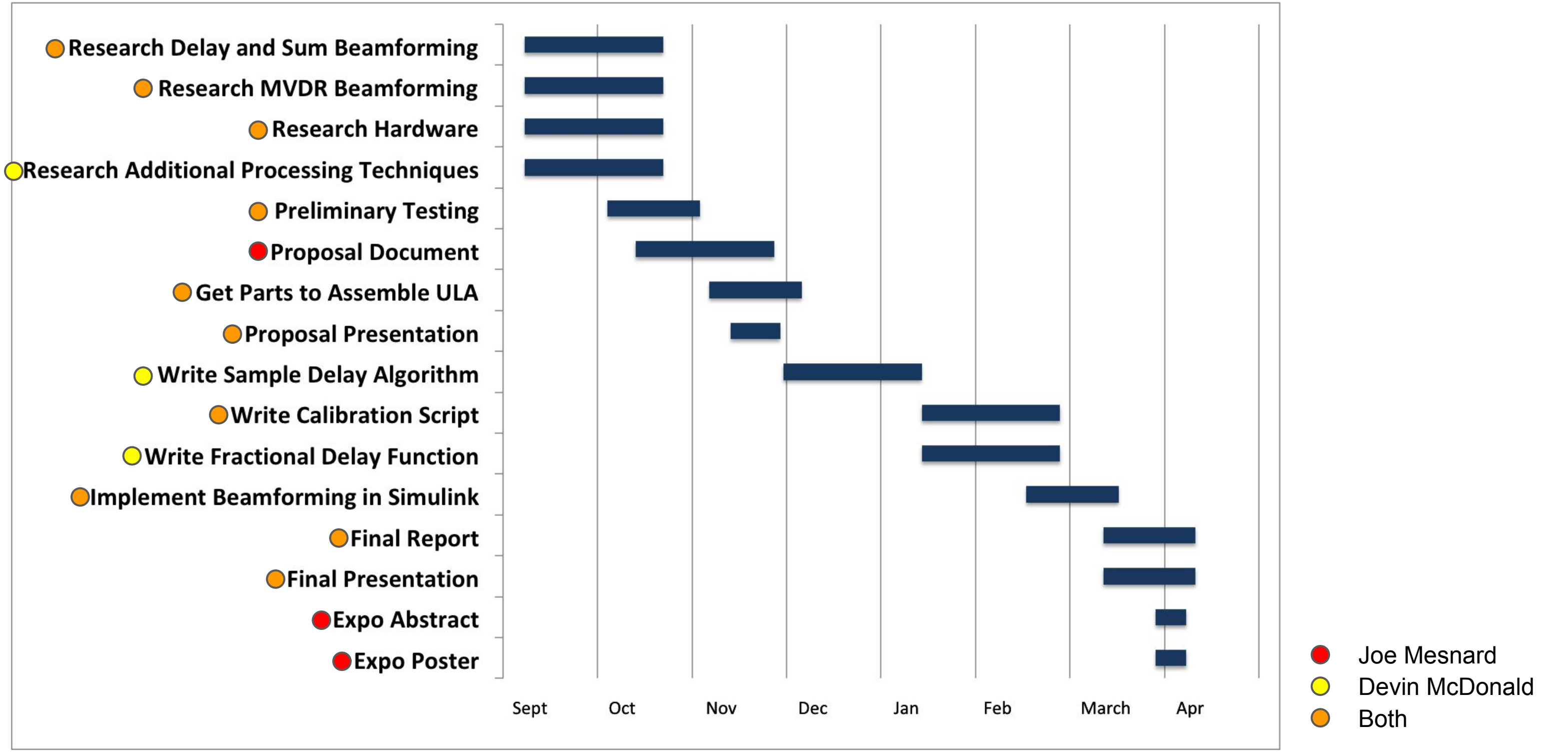


Figure 43 - Engineering efforts timeline

References

[1] "Texting and Driving Accident Statistics - Distracted Driving." *Edgarsnyder.com.* Accessed October 5, 2017. Available: https://www.edgarsnyder.com/car-accident/cause-of-accident/cell-phone/cell-phone-statistics.html

[2] "Phased Array System Toolbox - mvdrweights." (R2017b). *MathWorks.com.* Accessed July 14, 2017. Available: https://www.mathworks.com/help/phased/ref/mvdrweights.html

[3] "(Ultra) Cheap Microphone Array." *Maxime Ayotte*. Accessed November 28, 2017. Available: http://maximeayotte.wixsite.com/mypage/single-post/2015/06/25/Ultra-Cheap-microphone-array

[4] "Microphone Array Beamforming." *InvenSense.* Accessed November 28, 2017. Available: https://www.invensense.com/wp-content/uploads/2015/02/Microphone-Array-Beamforming.pdf

[5] "Delay Sum Beamforming." *The Lab Book Pages.* Accessed November 28, 2017. Available: http://www.labbookpages.co.uk/audio/beamforming/delaySum.html

Speech Intelligibility Enhancement using Microphone Array via Intra-Vehicular Beamforming

Devin McDonald, Joe Mesnard
Advisors: Dr. In Soo Ahn, Dr. Yufeng Lu

April 28th, 2018

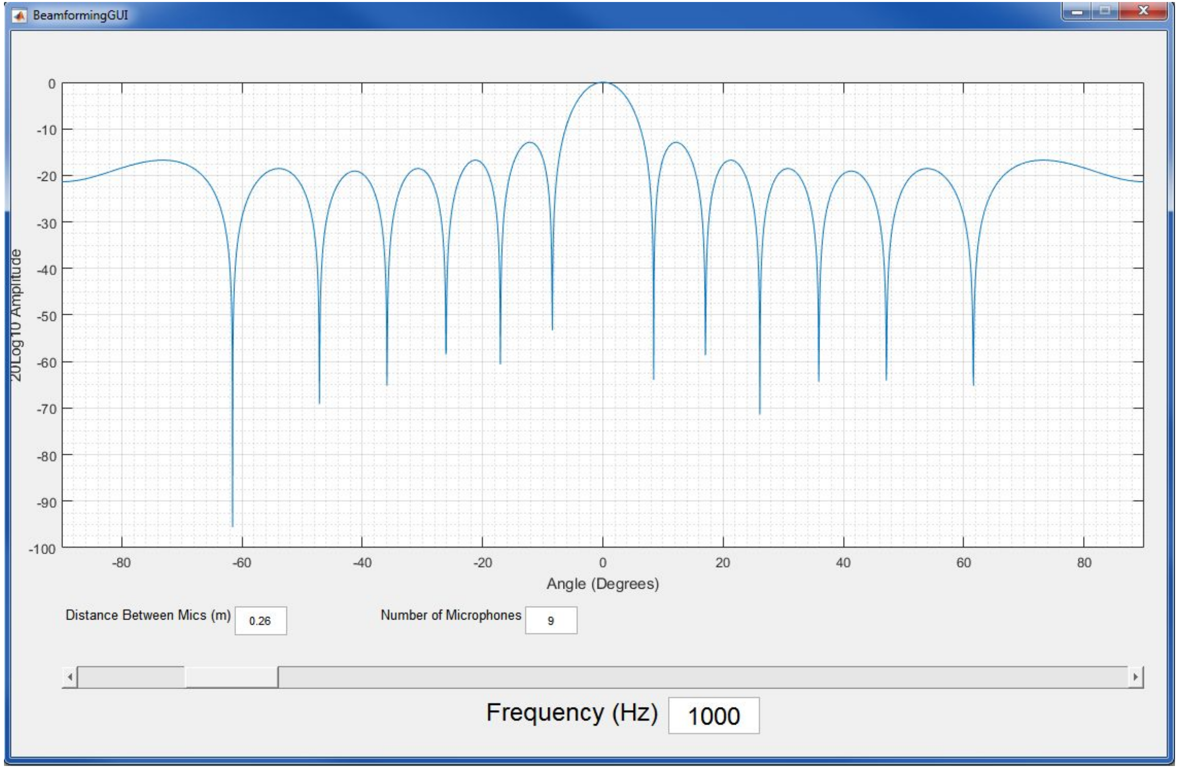
BRADLEY University

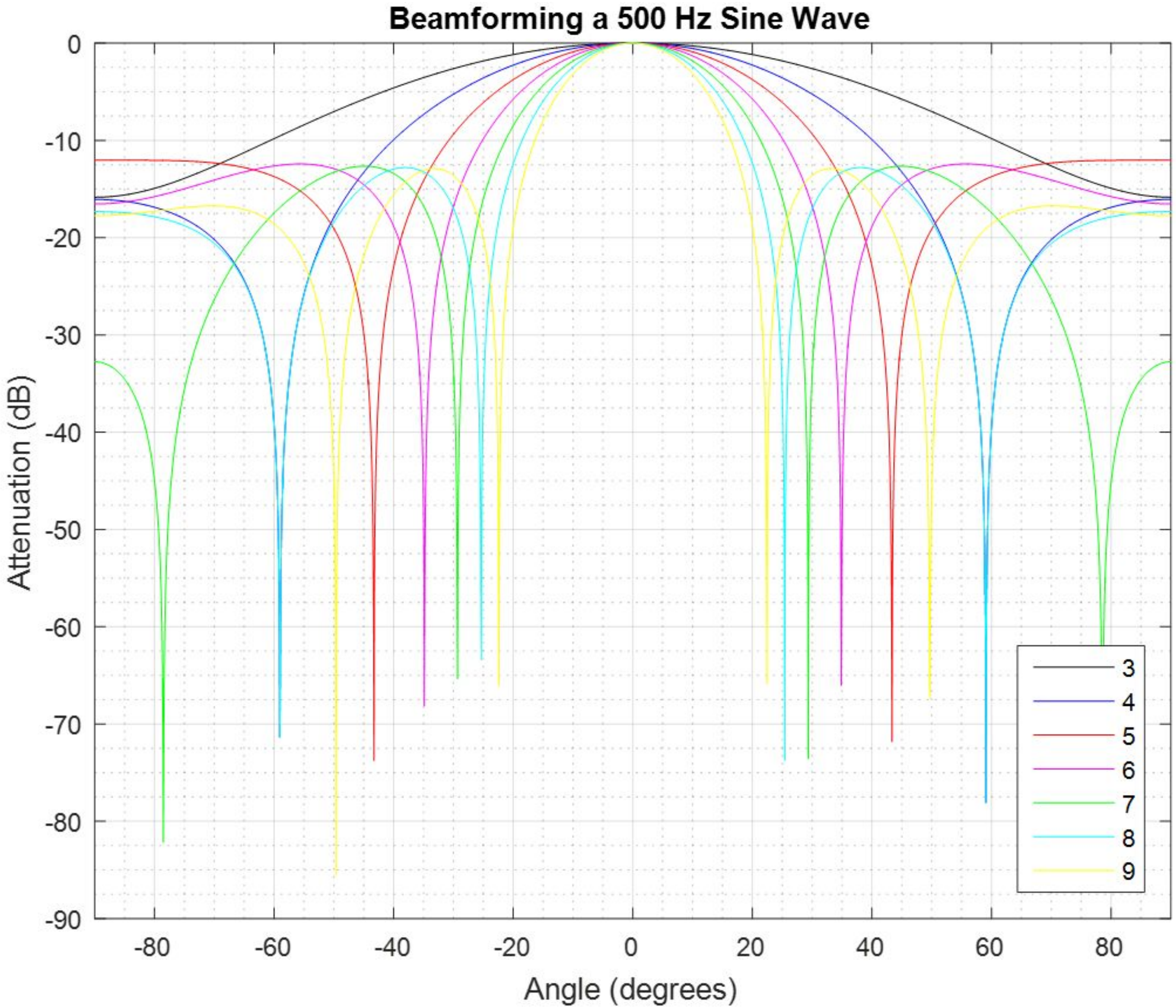
Appendix

Preliminary Results

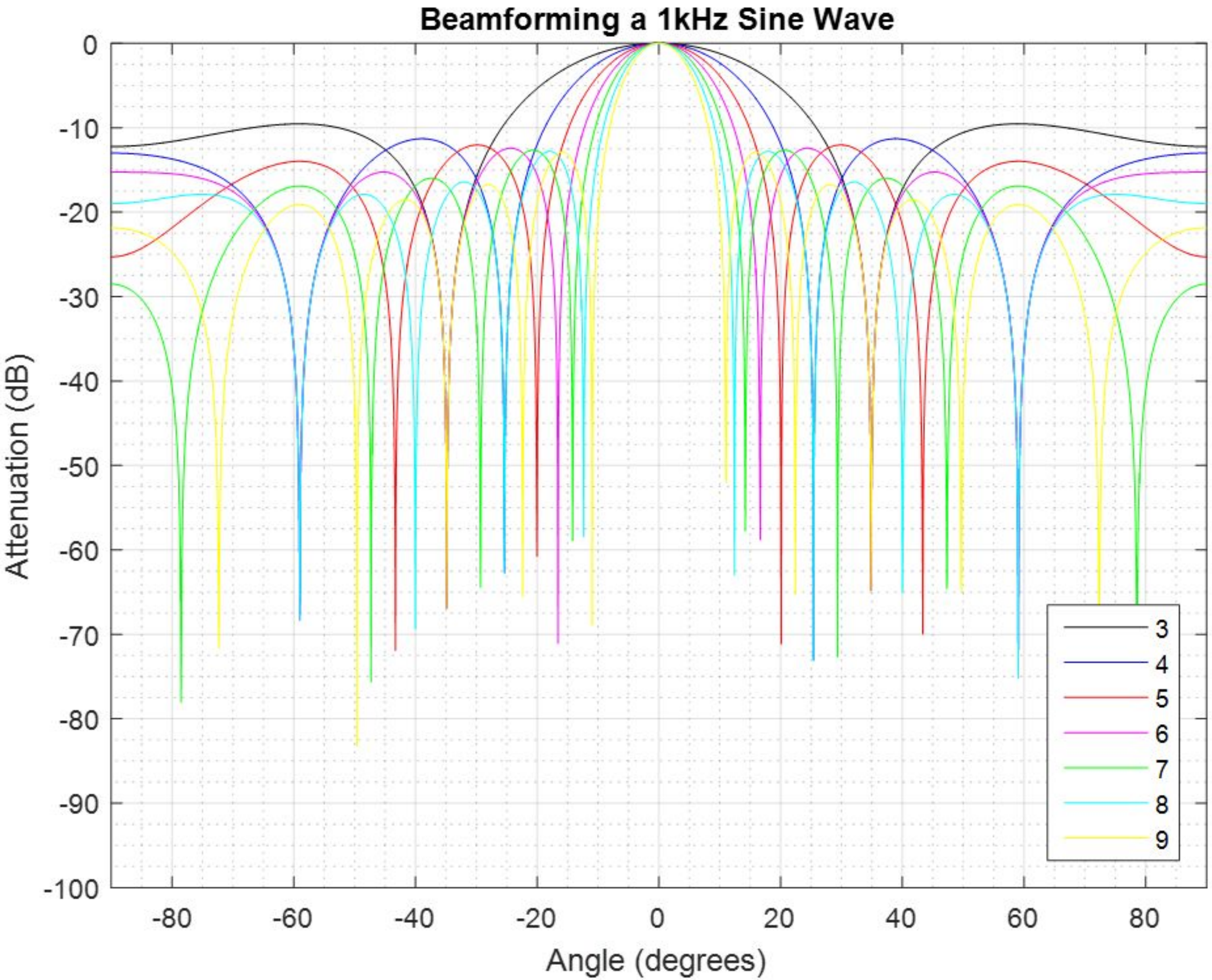
Second Test Setup

Matlab GUI for Beamforming

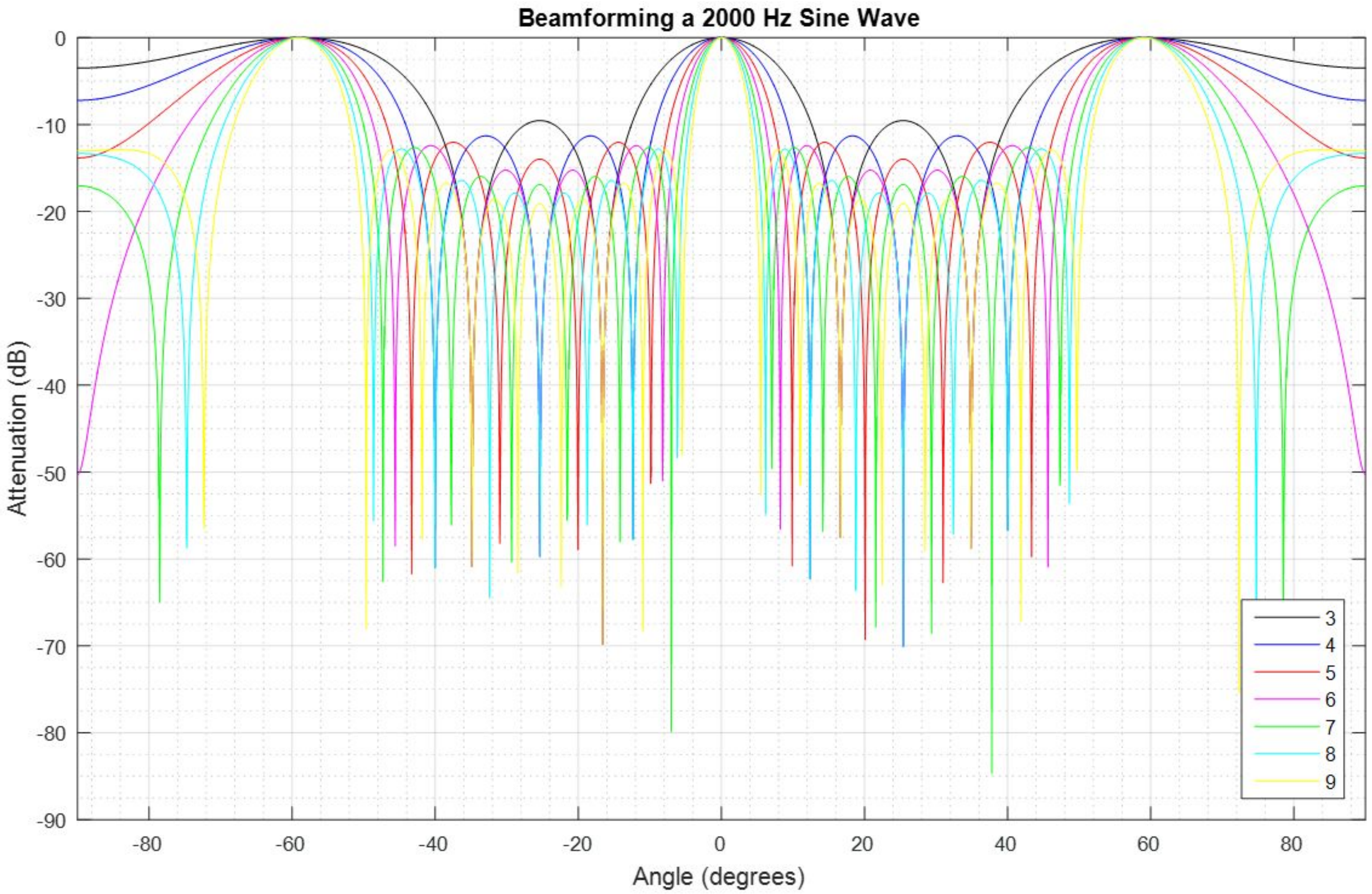




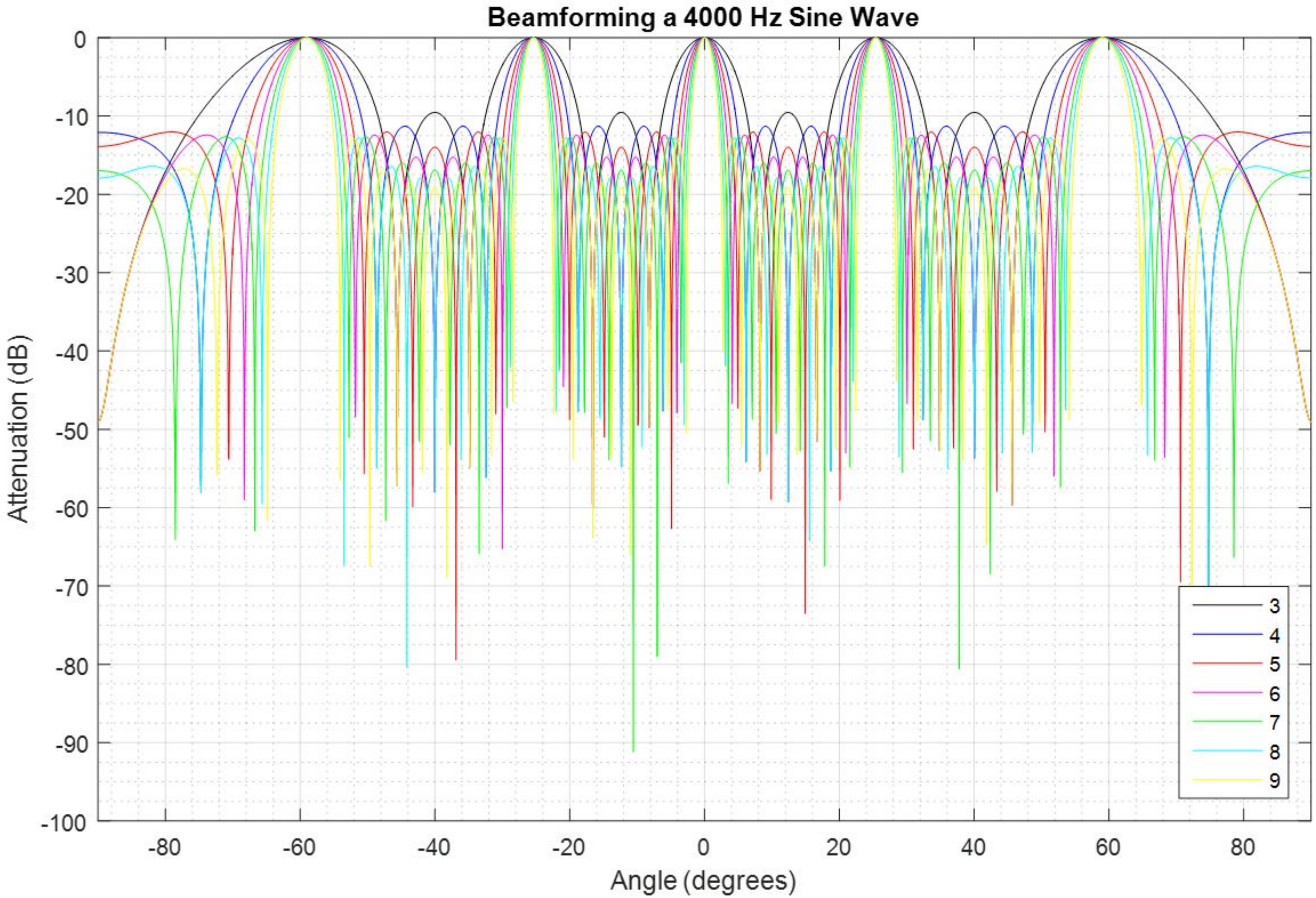
BRADLEY University



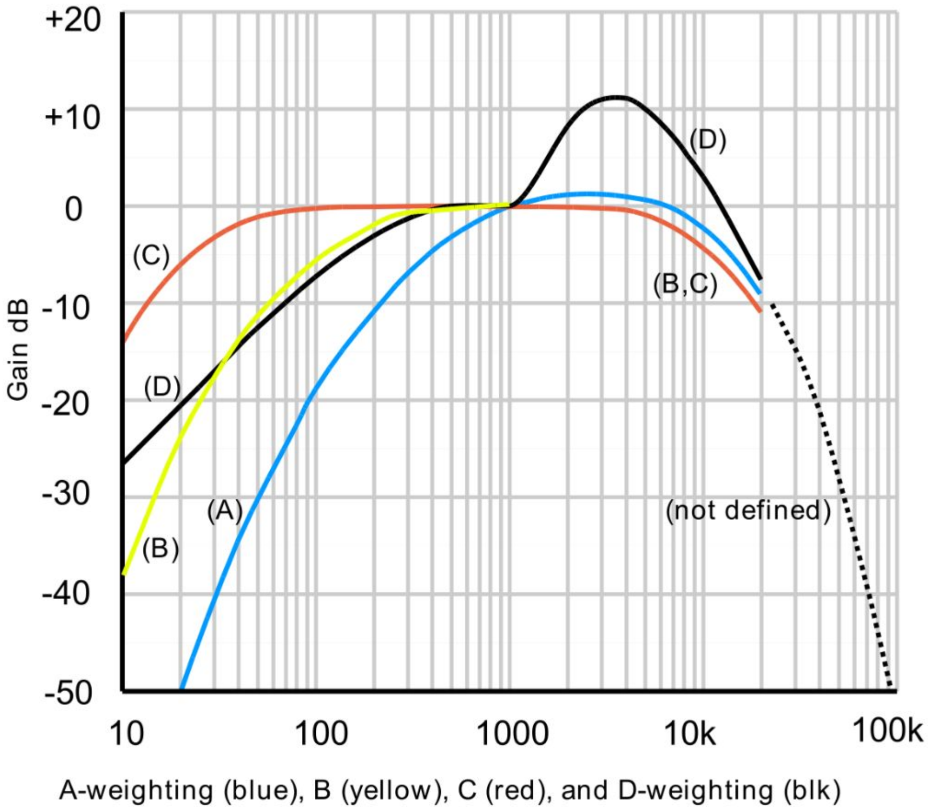
BRADLEY University



BRADLEY University



BRADLEY University



BRADLEY University

A-Weighting graph from https://en.wikipedia.org/wiki/A-weighting

Parts List With URLs

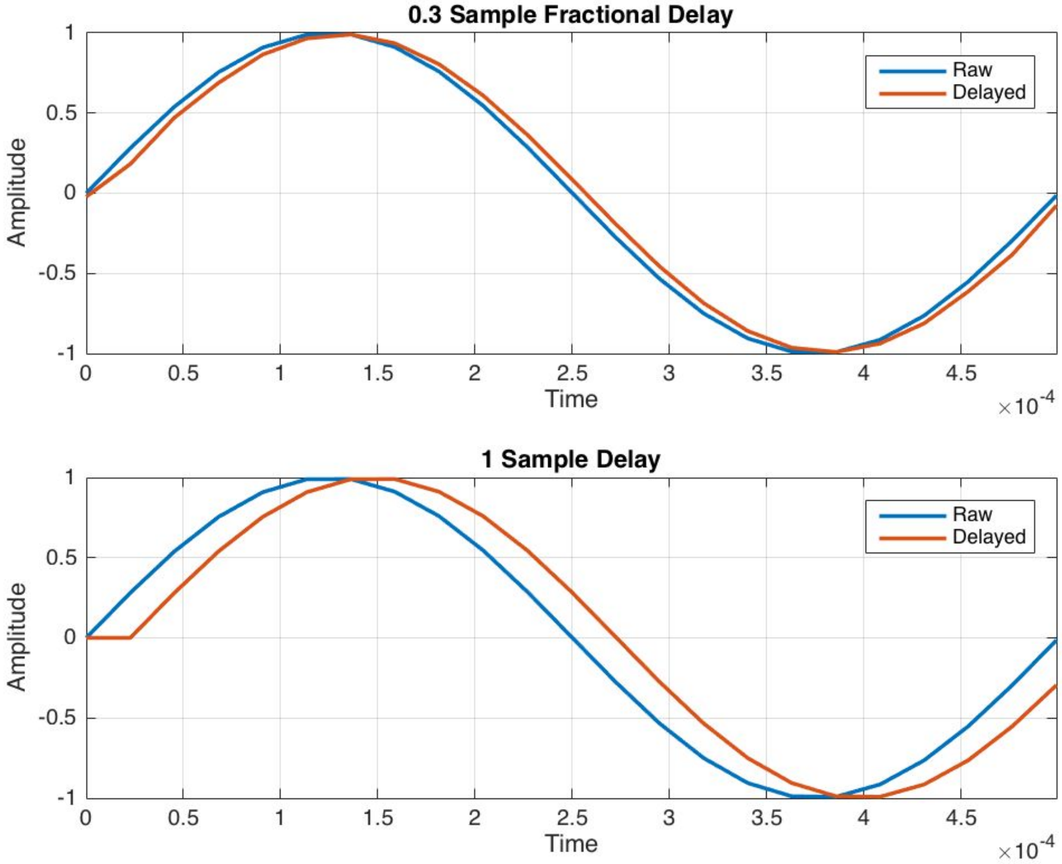
Quantity	Description	Price	Ext. Price
1	XLR Patch Cables https://www.amazon.com/Pack-Female-Microphone-Extension-Cable/dp/B01M0JQX2E/ref=sr_1_3?ie=UTF8&qid=1510258105&sr=8-3&keywords=3ft+xlr+pack&dpID=61YjshJDuwL&preST=_SY300_QL70_&dpSrc=srch	$31.75	$31.75
3	Behringer UltraVoice XM1800S Microphones https://www.amazon.com/Behringer-XM1800S-BEHRINGER-ULTRAVOICE/dp/B000NJ2TIE/ref=sr_1_4?ie=UTF8&qid=1510257881&sr=8-4&keywords=behringer+dynamic+microphone	$39.99	$119.97
5	Pro Black Adjustable Dual Plastic 2pcs Drum Microphone Clip https://www.amazon.com/Professional-Adjustable-Plastic-Microphone-Karaoke/dp/B06ZZCMJ26/ref=sr_1_87?s=musical-instruments&ie=UTF8&qid=1510262769&sr=1-87&keywords=mic+clamp	$7.44	$37.20
1	Scarlett 18i20 http://www.musiciansfriend.com/pro-audio/focusrite-scarlett-18i20-2nd-gen-usb-audio-interface/j35222000000000?cntry=us&source=3WWRWXGP&gclid=EAIaIQobChMIiu7F8a291wIV0LjACh36FQCZEAQYASABEgI3-_D_BwE&kwid=productads-adid^221957295827-device^c-plaid^323968843383-sku^J35222000000000@ADL4MF-adType^PLA	$499.99	$499.99

Fractional Delay

Fs = 44.1 kHz

f = 1 kHZ

Sampled sinc pulse



Demonstration of fractional delays [5]

BRADLEY University

Helpful Scales

Minimum Sample Delay at 44.1 kHz is 22.676 us

Time delay from a source 1 m away where microphones are 0.2 m apart is 57.737 us

The speed of sound is approximately 343 m/s

Wavelength of a 1 kHz signal is 0.343 m

System Description

N-Element Microphone Array

ULA of microphones will output signal via XLR.

Filters

A-Weighting Filters implemented in MATLAB/Simulink are designed to focus on the prominent frequencies of human speech (~500Hz to ~4kHz).

Delay

Delays will work as a part of the “Delay” and Sum beamforming algorithm

User input

The end user will be able to switch beam patterns to control where the beam is steered and who in the vehicle can be heard.

Audio Interface

The Focusrite Scarlett 18i20 will send digitized audio data from the microphones to the computer via USB.

Audio System Toolbox

The audio system toolbox in Simulink will be used to communicate with the audio interface and get stream data into Simulink.