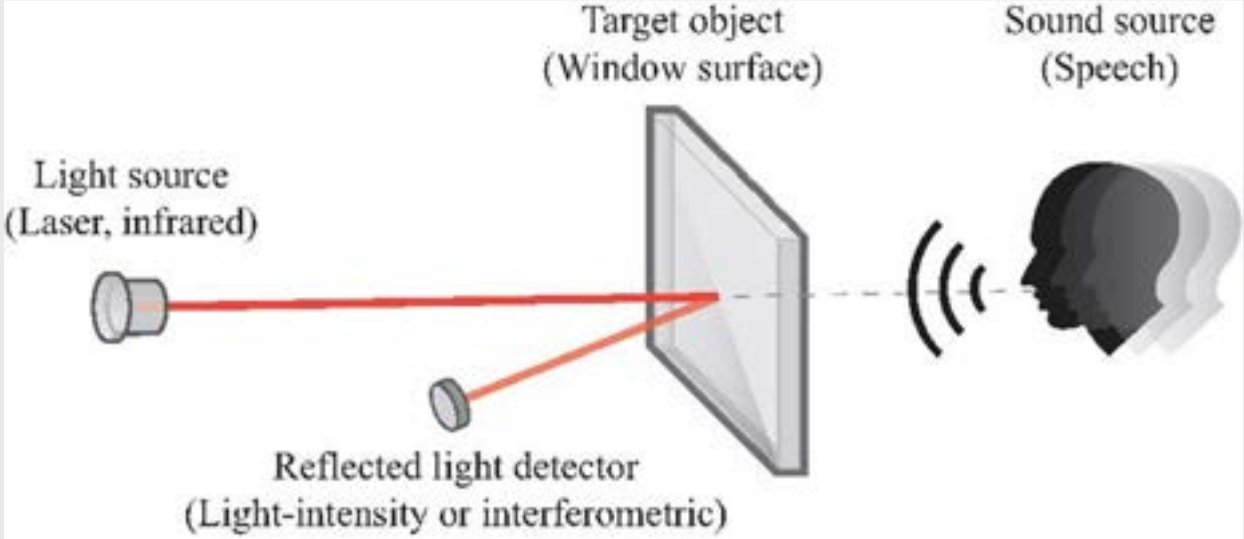
MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

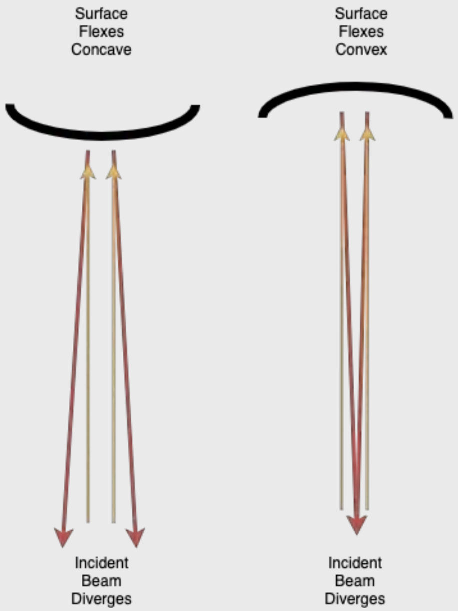
# Laser Microphone Methods

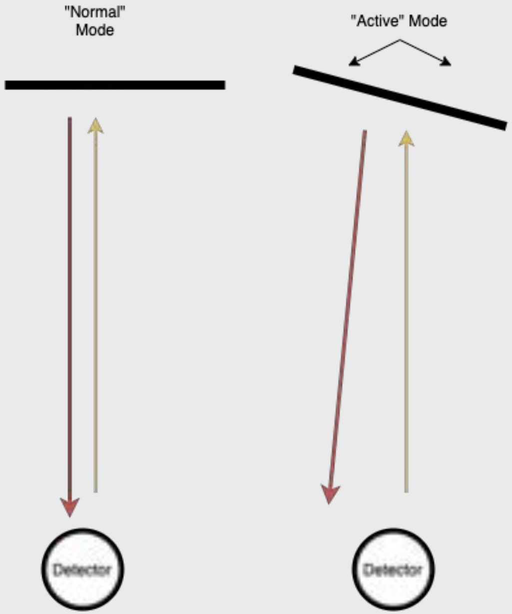
EECS 452: Andrew Sager, Ryan Aridi, Evan Arora, Kyle Liebler

# Background


Target object
(Window surface)
Sound source
(Speech)
Light source
(Laser, infrared)
Reflected light detector
(Light-intensity or interferometric)

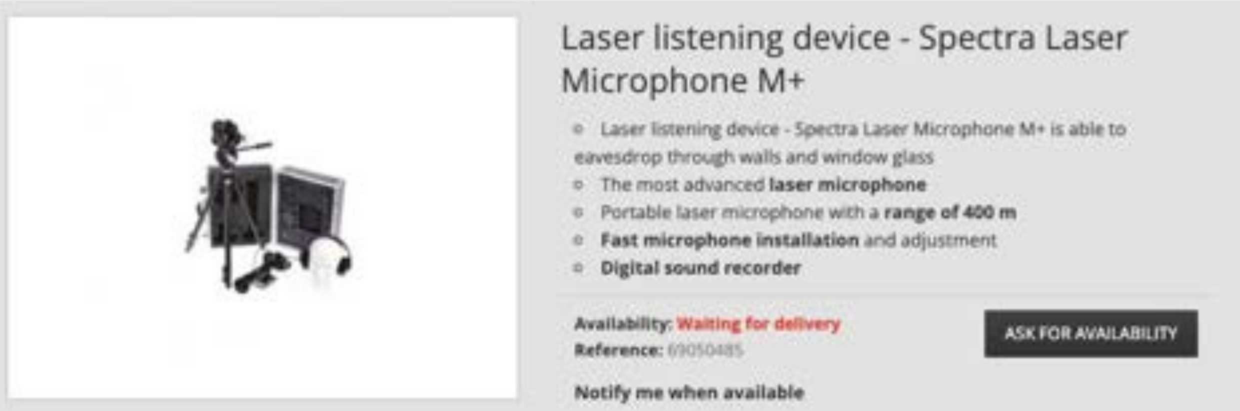
# How it Works


Surface Flexes Concave
Surface Flexes Convex
Incident Beam Diverges
Incident Beam Diverges


"Normal" Mode
"Active" Mode
Detector
Detector

# Previous Work

Pioneered by Soviet Union - Buran eavesdropping system 1947 - Léon Theremin

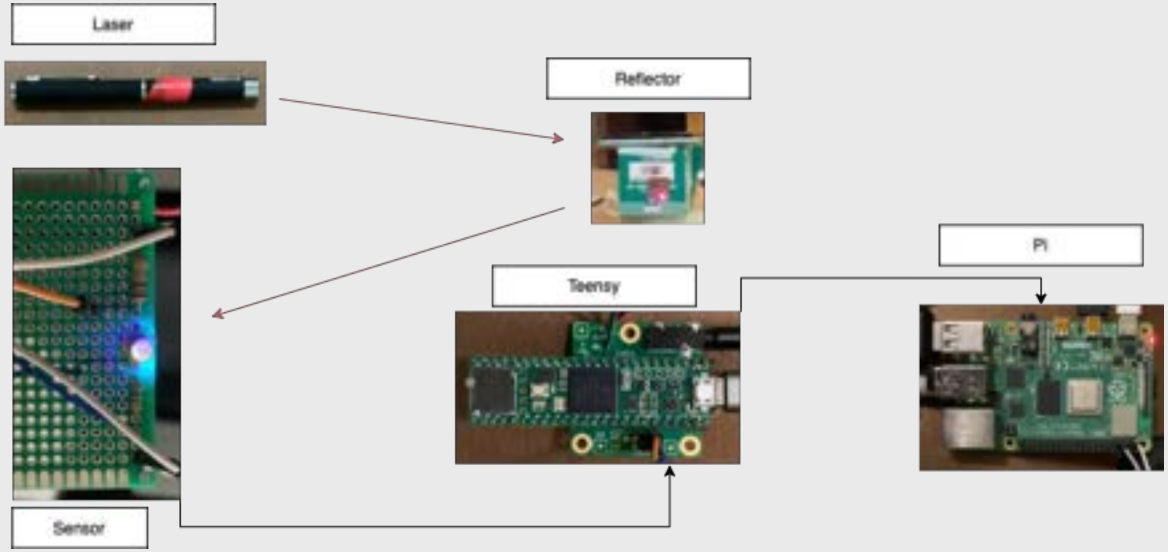



DetectiveStore.net
$60,000

# Our Goals

- Recover intelligible speech using inexpensive and simple hardware setup.
- Investigate DSP methods to reduce noise and enhance signal.
- “Does the audio sound better post processing?”

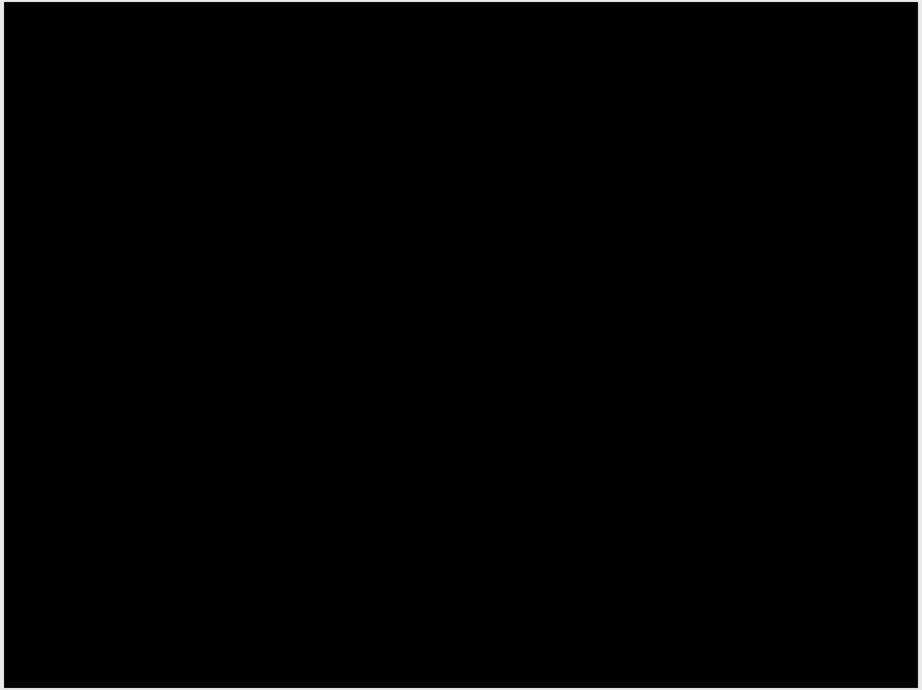
# What We Built




Bathroom Mirror

Cavity - Resonator

# Demo



# 4 (Primary) Algorithms

## Filtering - Teensy

Equiripple FIR Bandpass filter with cutoff frequencies of 100 Hz and 4.5 kHz.

## Spectral Subtraction - Pi

Spectral magnitude reduced by the magnitude of an earlier recorded noise signal (no speech present).

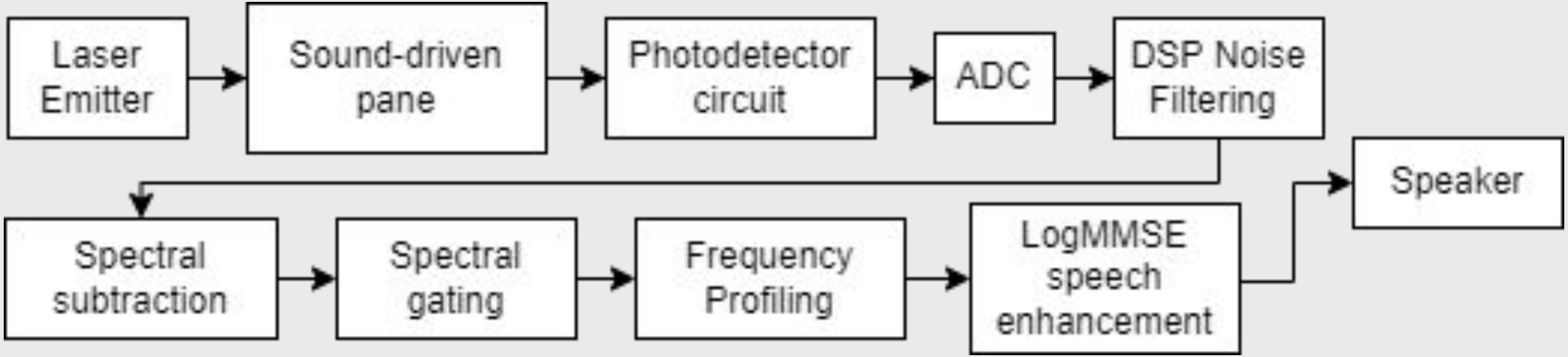
## Spectral Gating - Pi

Below and above a statistically set number of standard deviations from the sample’s average spectral magnitude, everything is zeroed.

## Mirror Profiling - Pi

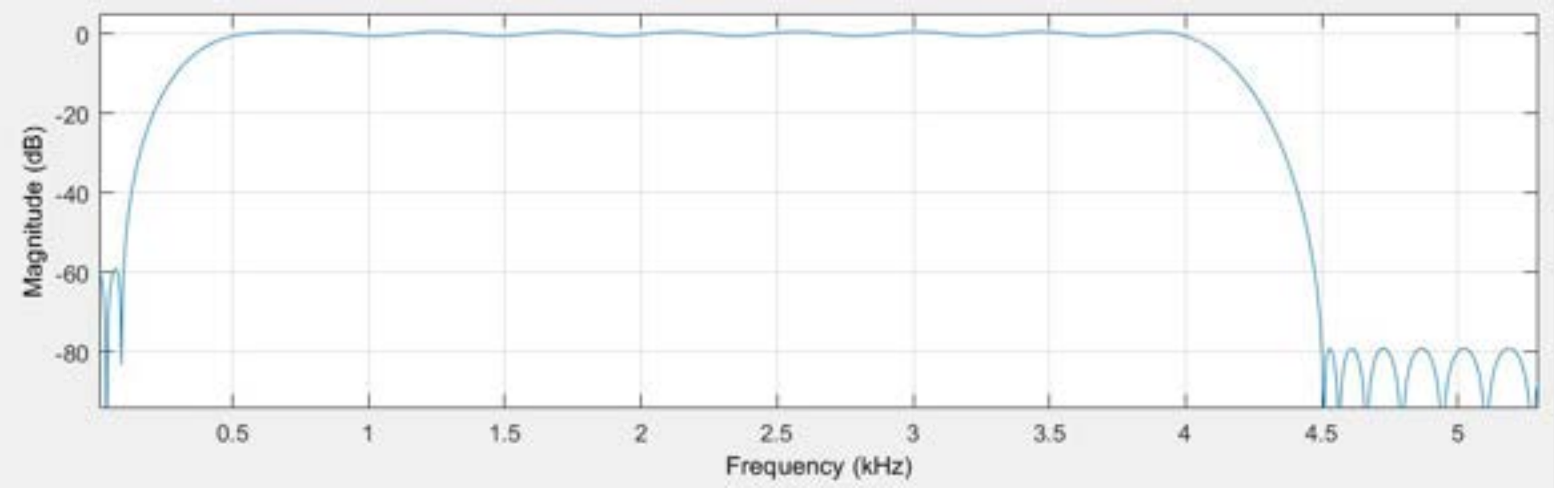
Account for frequency response of reflecting surface

# Signal Flow Chart

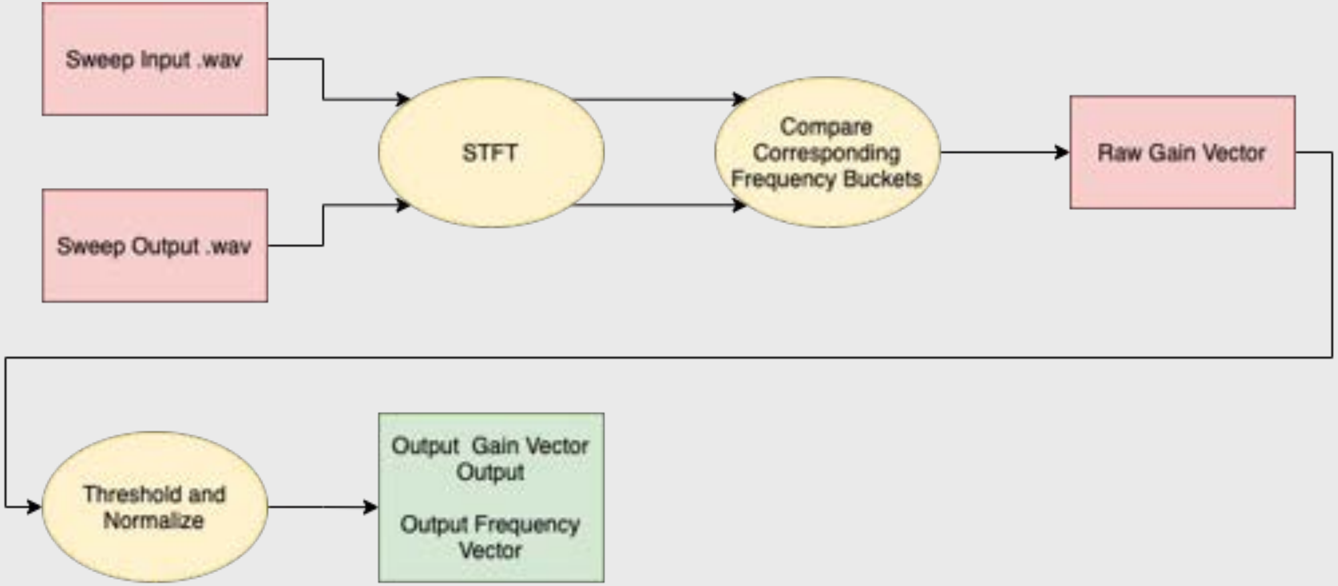



Test Combinations

# Bare Metal Filtering



# Mirror Profiling

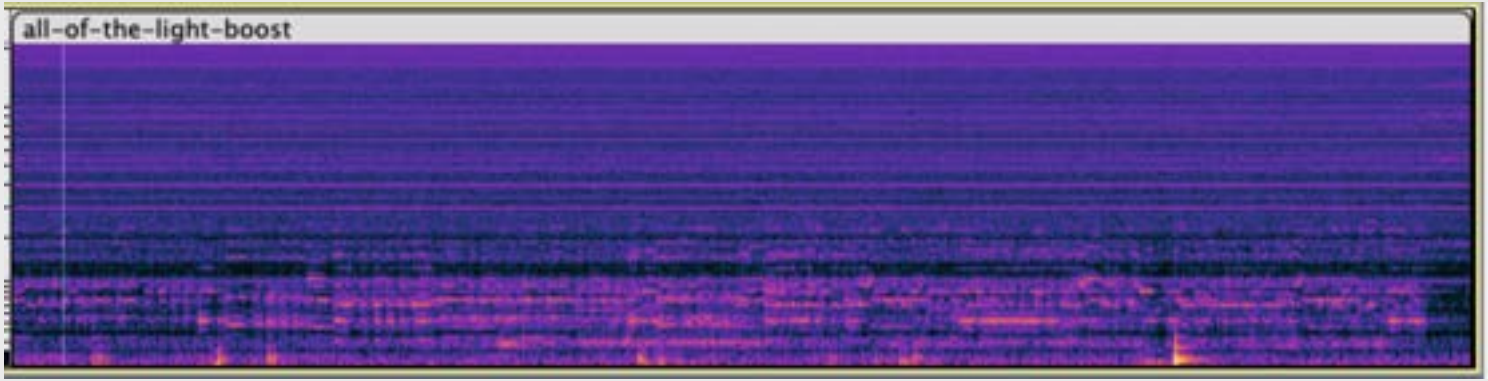


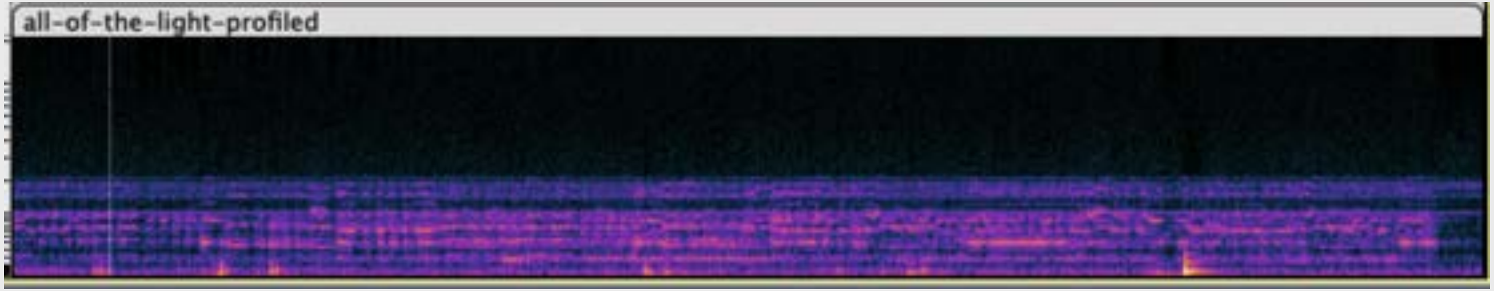
# Mirror Profiling

## 50-500 Hz Example


At some frequencies, output is **severely** attenuated (less than noise). I exclude these frequencies during gain adjustment

# Mirror Profiling



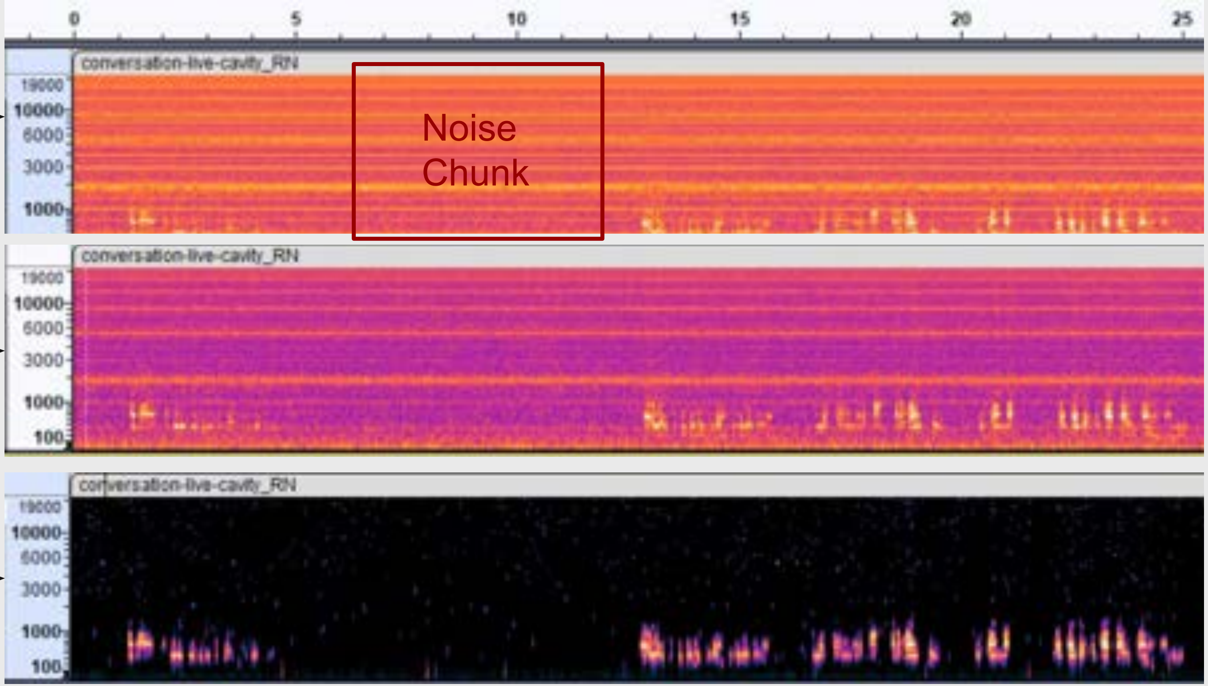
MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

# Noise Removal - Spectral Subtraction

**Raw Signal**

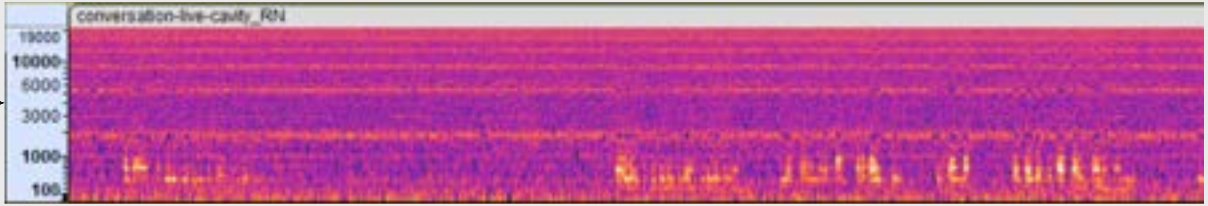
**1x Spectral Subtraction**

**100x Spectral Subtraction**

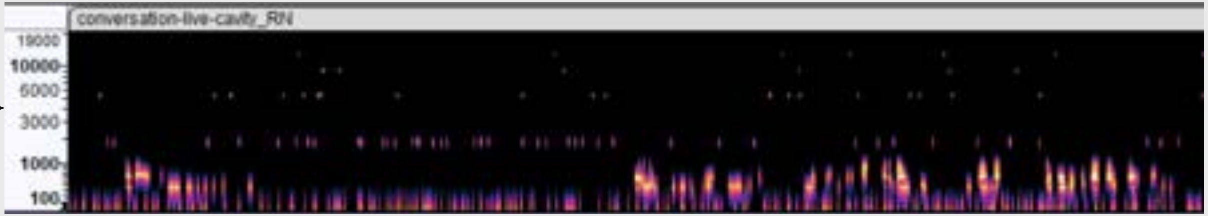


# Spectral Subtraction & Manual Gating

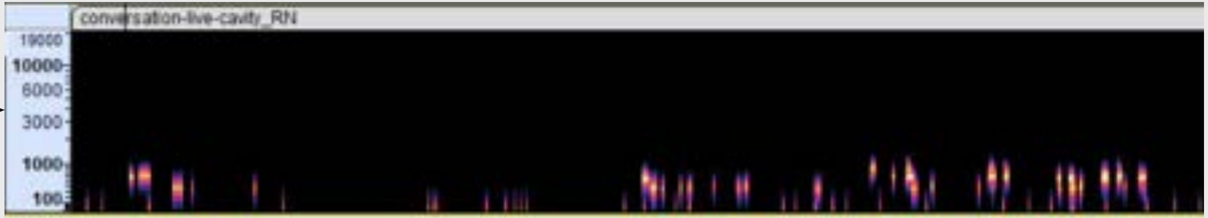
**0 σ Gating**




**5 σ Gating**

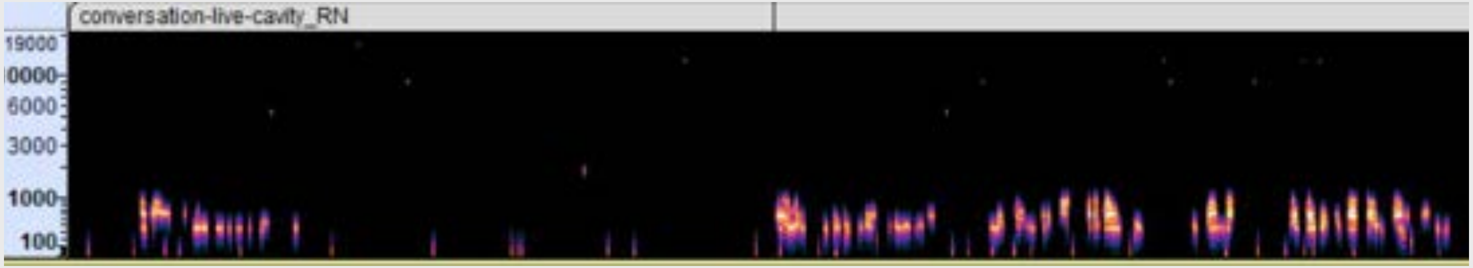



**10 σ Gating**

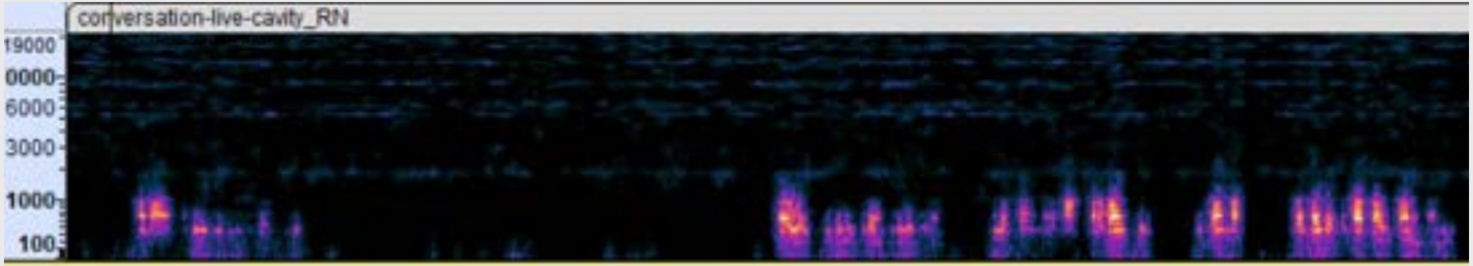


# Manual vs Statistical Gating

**3x Subtraction at 1σ Gating (~0.5 sec)**

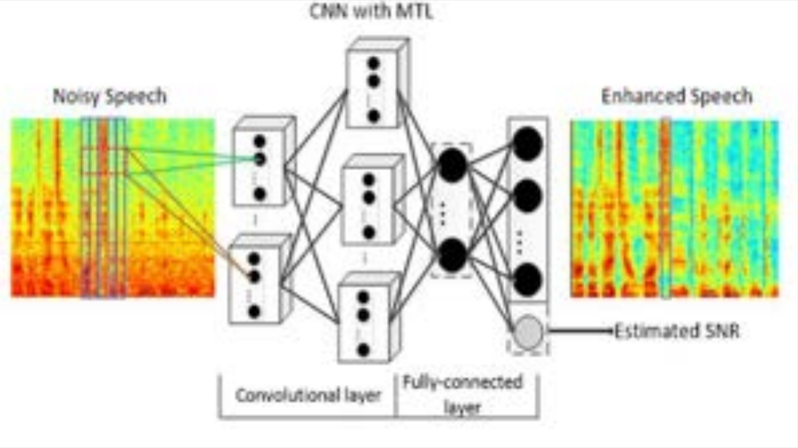



**1x Subtraction with Statistical Gating (~0.5 sec)**

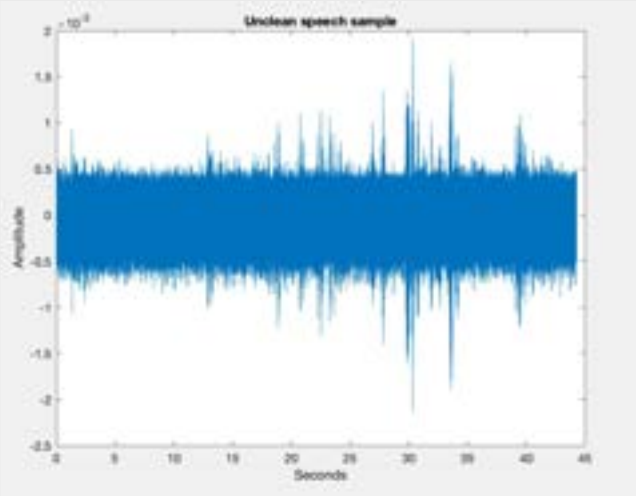


# Speech Enhancement

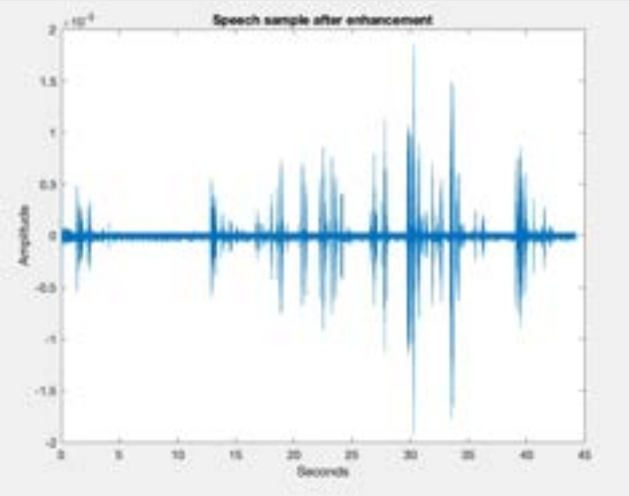
- **Machine Learning**
  - Many modern speech enhancement techniques utilize neural networks to classify audio
  - For a variety of reasons, machine learning was not feasible for a project of this scale



# Speech Enhancement




→




MMSE log-STSA estimator

# Testing Scheme

- Tested algorithms on  four unique inputs
  - Calibration input
  - Live conversations
  - Pre-recorded conversations
  - Music
- Two reflecting mediums:
  - Custom-built resonator cavity
  - Bathroom mirror

# Testing Scheme

We evaluate the success of each algorithm, and combinations of algorithms on two qualitative measures: noise and speech intelligibility.

We randomized and blindly scored 25 samples. In total, around ~30 minutes of audio

# Results

| **Method / Test** | A | B | C | D | E |
|---|---|---|---|---|---|
| *Band Pass Filtering* | | | | | |
| *Spectral Methods* | | | | | |
| *Mirror Profiling* | | | | | |
| *Speech Enhancement* | | | | | |
| **Noise** | 1.50 | 2.55 | 2.42 | 2.32 | 2.27 |
| **Speech Intelligibility** | 2.12 | 2.22 | 2.22 | 2.00 | 2.10 |
| **Average Score** | 1.81 | **2.39** | **2.33** | 2.16 | 2.19 |

Fig 7: Testing Results Synopsis

# Software Issues

- Noise removal requires tuning
- Implementing noise removal on an audio stream led to some issues with data types and buffer size
- Additional speech enhancement strategies had overlap with our spectral subtraction algorithm

# Hardware Issues

- Laser Tweaks
  - Photodiodes were much noisier than phototransistors
  - Trying to build a laser driver was an unnecessary complication
  - Resultant signal was much better sometimes when reflection as slightly offset from the transistor rather than directly incident to it
- Raspberry Pi Power Issues
  - With so many devices connected to the Pi, we maxed out the power output

# Even more Issues

- Speech to Text
  - Were not successful enough to use this as a measure
- Inconsistency in Raw signal Quality
  - Between similar trials, huge differences (too setup dependent)
- Lack of sensitivity
  - Forget about whispering or double paned windows.

# Future Work

- It is possible that using a second phototransistor to capture environmental noise could improve the quality of our resultant signal
- Our final testbench tested the project indoors and at relatively short (<25 ft) distances. More testing can be done in outdoor environments
- More testing can also be done with a variety of reflective surfaces (window, drywall, house plant leaf)
- Add a feedback loop to tune the noise removal algorithm parameters

# Future Work

- Techniques used in speech dereverberation might be applicable here and could cause large improvements in speech intelligibility.
  - They were not pursued due to their necessity for more than one audio input channel or deep neural network methods not feasible for implementation on a Raspberry Pi (at least with our current skill-sets).

# Acknowledgements

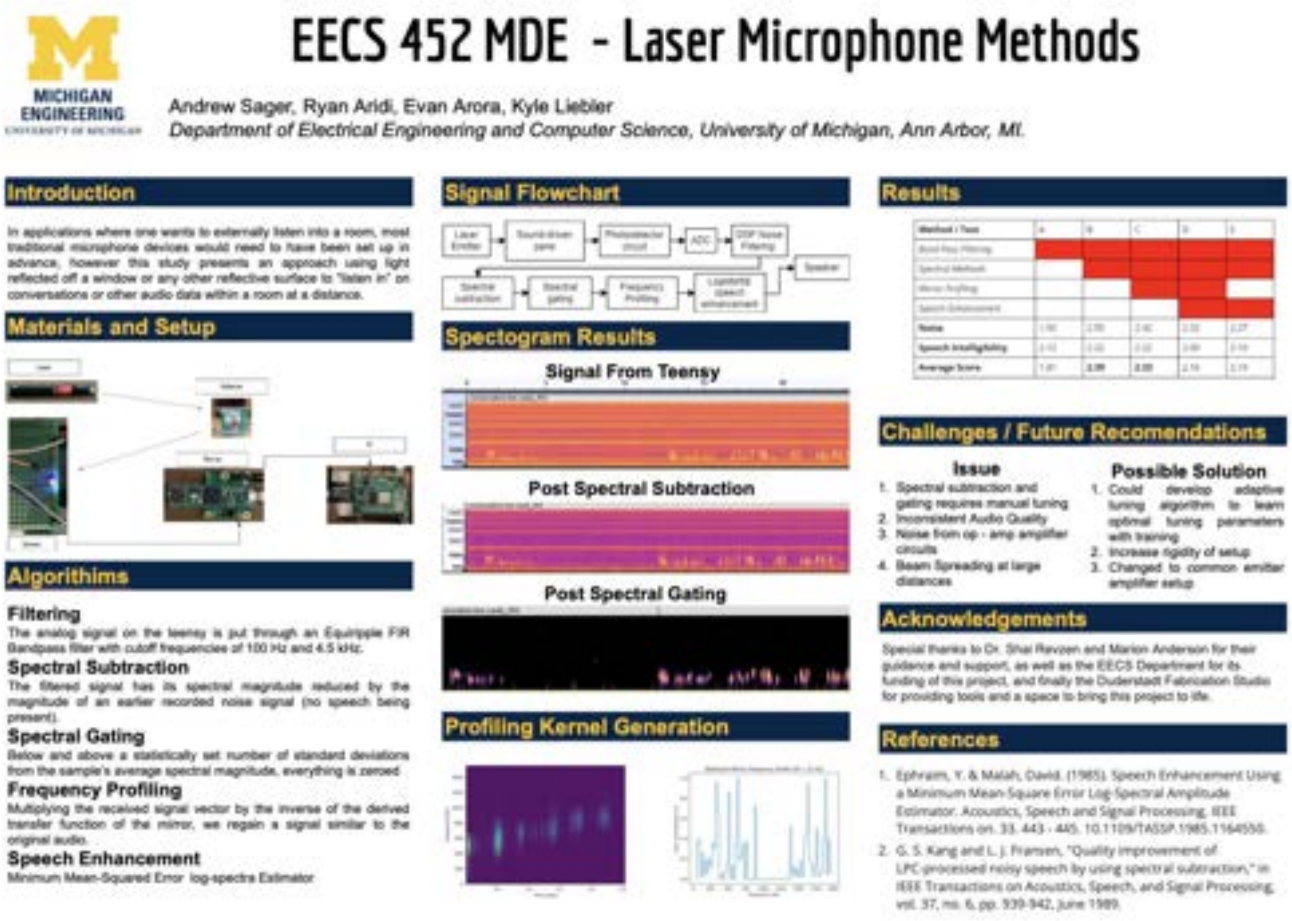
We would like to thank:

- Professor Shai Revzen
- Marion Anderson
- EECS Department
- Duderstadt Fabrication Studio

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

# Thank you!

Link


MICHIGAN ENGINEERING
EECS 452 MDE - Laser Microphone Methods
Andrew Sager, Ryan Aridi, Evan Arora, Kyle Liebler
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI
Introduction
In applications where one wants to externally listen into a room, most traditional microphone devices would need to have been set up in advance, however this study presents an approach using light reflected off a window or any other reflective surface to "listen in" on conversations or other audio data within a room at a distance.
Materials and Setup
Algorithims
Filtering
The analog signal on the teensy is put through an Equiripple FIR Bandpass filter with cutoff frequencies of 100 Hz and 4.5 kHz.
Spectral Subtraction
The filtered signal has its spectral magnitude reduced by the magnitude of an earlier recorded noise signal (no speech being present).
Spectral Gating
Below and above a statistically set number of standard deviations from the sample's average spectral magnitude, everything is zeroed
Frequency Profiling
Multiplying the received signal vector by the inverse of the derived transfer function of the mirror, we regain a signal similar to the original audio.
Speech Enhancement
Minimum Mean-Squared Error log-spectra Estimator
Signal Flowchart
Spectogram Results
Signal From Teensy
Post Spectral Subtraction
Post Spectral Gating
Profiling Kernel Generation
Results
Challenges / Future Recomendations
Issue
1. Spectral subtraction and gating requires manual tuning
2. Inconsistent Audio Quality
3. Noise from op - amp amplifier circuits
4. Beam Spreading at large distances
Possible Solution
1. Could develop adaptive tuning algorithm to learn optimal tuning parameters with training
2. Increase rigidity of setup
3. Changed to common emitter amplifier setup
Acknowledgements
Special thanks to Dr. Shai Revzen and Marion Anderson for their guidance and support, as well as the EECS Department for its funding of this project, and finally the Duderstadt Fabrication Studio for providing tools and a space to bring this project to life.
References
1. Ephraim, Y. & Malah, David. (1985). Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator. Acoustics, Speech and Signal Processing, IEEE Transactions on. 33. 443 - 445. 10.1109/TASSP.1985.1164550.
2. G. S. Kang and L. J. Fransen, "Quality improvement of LPC-processed noisy speech by using spectral subtraction," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 6, pp. 939-942, June 1989.