Cyber@Ben-Gurion University of the Negev | Israel National Cyber Bureau

Cyber Security Research Center

BadVibes

Sound Recovery Using Bulbs' Vibrations

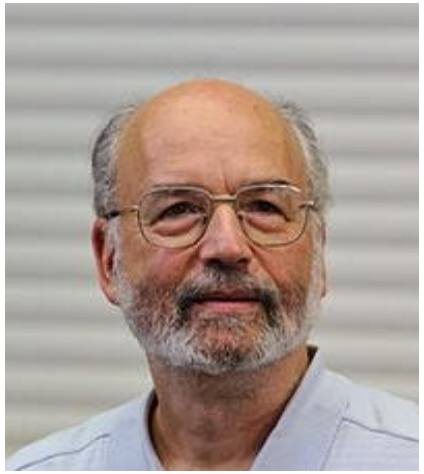
Dr. Boris Zadov and Yaron Pirutin

Ben-Gurion University of the Negev

Researchers

Ben Nassi

Yaron Pirutin



Prof. Adi Shamir

Prof. Yuval Elovici

Dr. Boris Zadov

Ben-Gurion University of the Negev

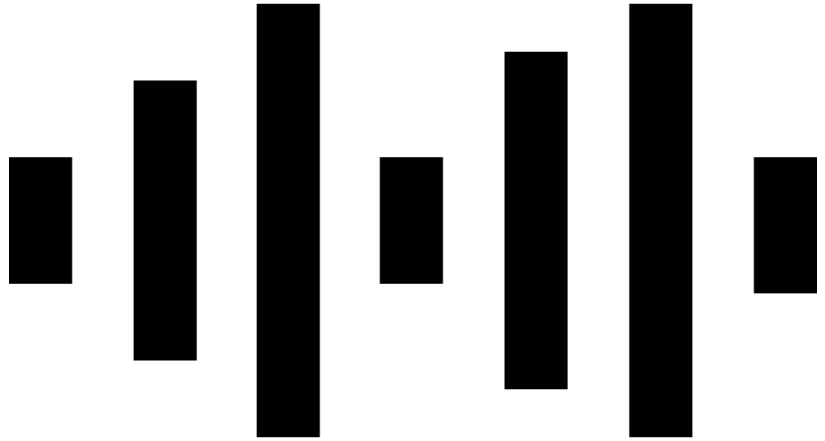
מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Agenda

- Background
  - What is sound?
  - How sound recording works?
  - Related Eavesdropping research
- BadVibes's Threat Model
- Physical Analysis
- Algorithm
- Evaluation
  - Testing the attack
- Potential Improvements
- Q&A

Background

Sound Wave



1. A sound wave is air traveling through space.
2. Humans hearing range is from 20 Hz to 20 kHz.
3. Human speech 150-4300 Hz
4. Speech can be recognized at 150-1000 Hz

How Does Sound Recording Work?

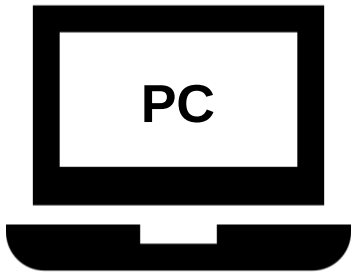
- When a sound wave hits the surface of an object, it causes the surface to vibrate slightly.
- These vibrations cannot be seen by the human eye.

- In sound recording, microphones are used to convert sound waves to electrical signals.

How Does Sound Recording Work?

Microphone contains three primary components:

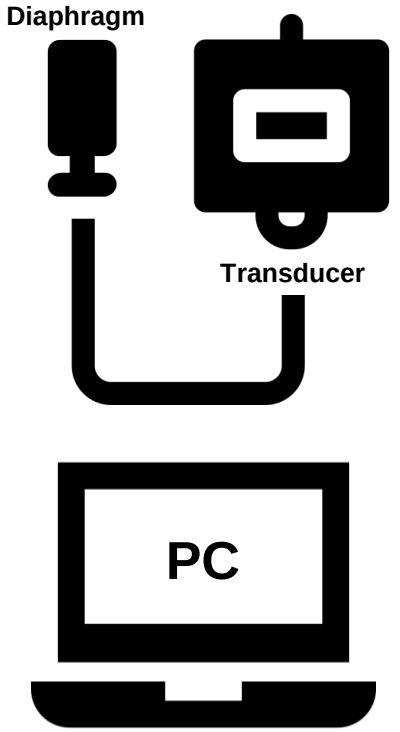
1. Diaphragm: a thin piece of material (e.g., plastic, aluminum) that vibrates when it is struck by sound waves.



How Does Sound Recording Work?

Microphone contains three primary components:

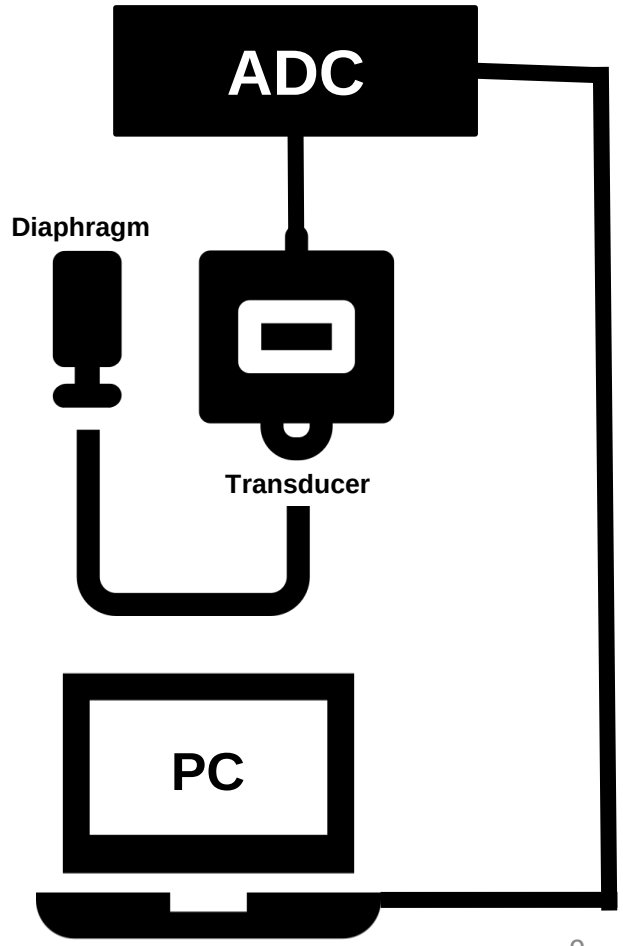
1. Diaphragm: a thin piece of material (e.g., plastic, aluminum) that vibrates when it is struck by sound waves.
2. Transducer: used to convert the diaphragm's vibrations to analog electrical signal.



How Does Sound Recording Work?

Microphone contains three primary components:

1. Diaphragm: a thin piece of material (e.g., plastic, aluminum) that vibrates when it is struck by sound waves.
2. Transducer: used to convert the diaphragm's vibrations to analog electrical signal.
3. ADC (analog-to-digital converter): used to convert the analog electric signal to digital data at standard audio sample rates (e.g., 44.1 kHz).

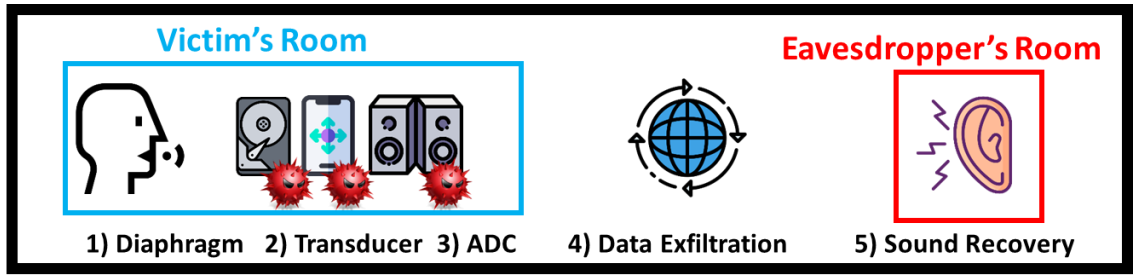


Related Work

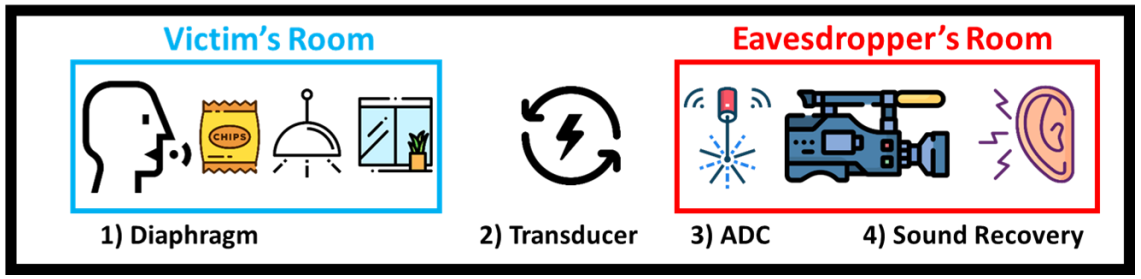
Related Eavesdropping Research

- In recent years, various eavesdropping methods have been developed by the scientific community. These methods can be divided into two main categories:

Internal Methods

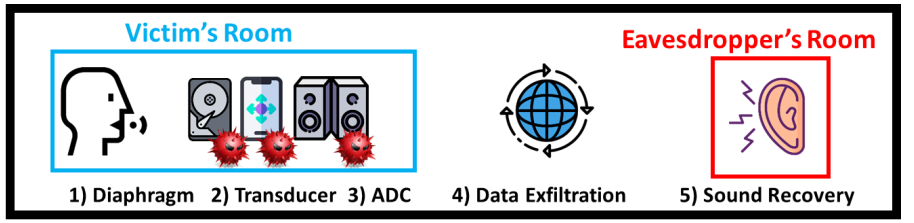


External Methods



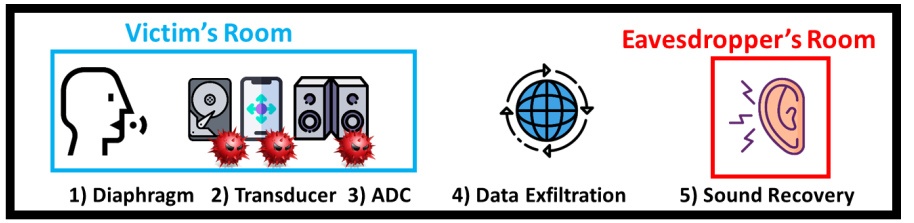
Internal Methods

- Methods which obtain data by using a device located in proximity to a target



Internal Methods

- Methods which obtain data by using a device located in proximity to a target



Motion Sensors

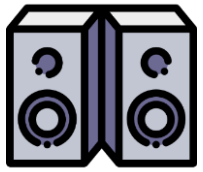
Gyroscope [1]

Accelerometer [2-4]

Output Devices

Speakers [5]

Vibration Motor [6]

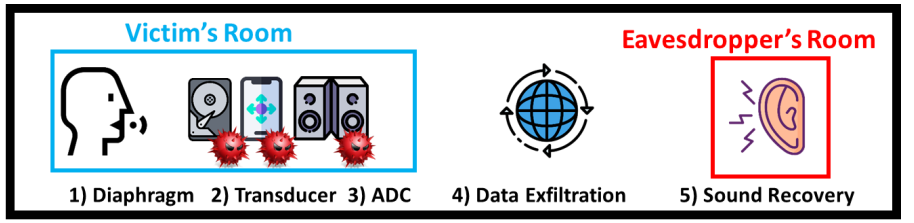


Misc.

Hard Drive [7]

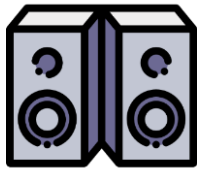
Internal Methods

➢ Methods which obtain data by using a device located in proximity to a target



Motion Sensors
Gyroscope [1]
Accelerometer [2-4]

Output Devices
Speakers [5]
Vibration Motor [6]



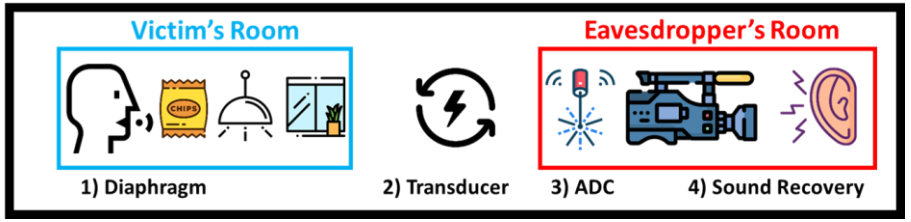
Misc.
Hard Drive [7]

Limitation:

- Internal methods require the attacker to **place a device near a target** in order to obtain and exfiltrate data

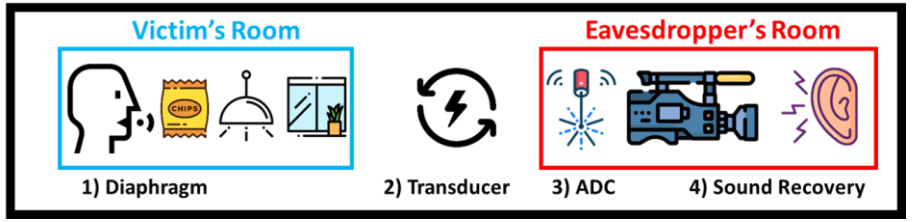
External Methods

- Methods that rely on data obtained by a device that is not located near a victim (individuals or organizations)



External Methods

➢ Methods that rely on data obtained by a device that is not located near a victim (individuals or organizations)



Laser Microphone [8]

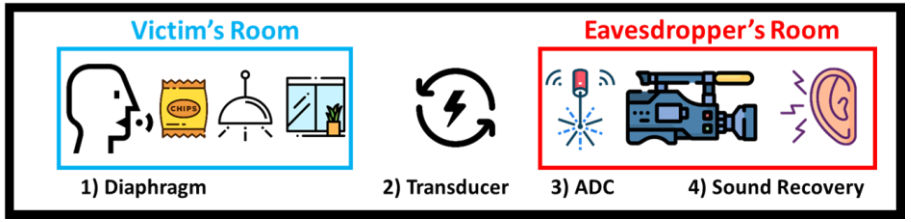
Uses a laser transceiver to recover sound by directing a laser beam at an object and analyzing its vibrations as a response to sound.

Limitation:

- It is **active** – the laser beam can be detected by victims/ organizations by using an optical sensor.

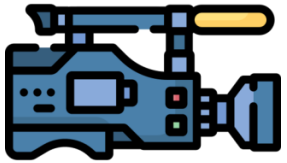
External Methods

➢ Methods that rely on data obtained by a device that is not located near a victim



Visual Microphone [9]

Uses a high frequency video camera (~2000 FPS) to recover sound by analyzing the object's (e.g., a bag of chips) vibrations as a response to sound.



Limitation:

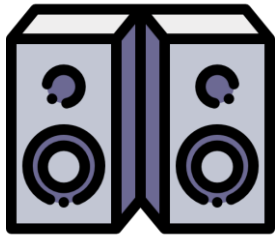
- It cannot be applied in **real time** – it takes a few hours to reconstruct sound.
- It is expensive – it requires **expensive** equipment and heavy computational resources.

Limitations of Existing Methods

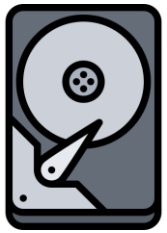
Internal Methods

Motion Sensors

Output Devices



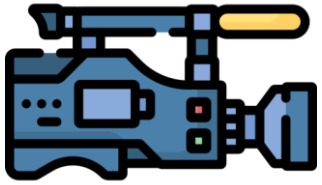
Misc.



External Methods

Laser Microphone

Visual Microphone



Each method is limited by one of the following:

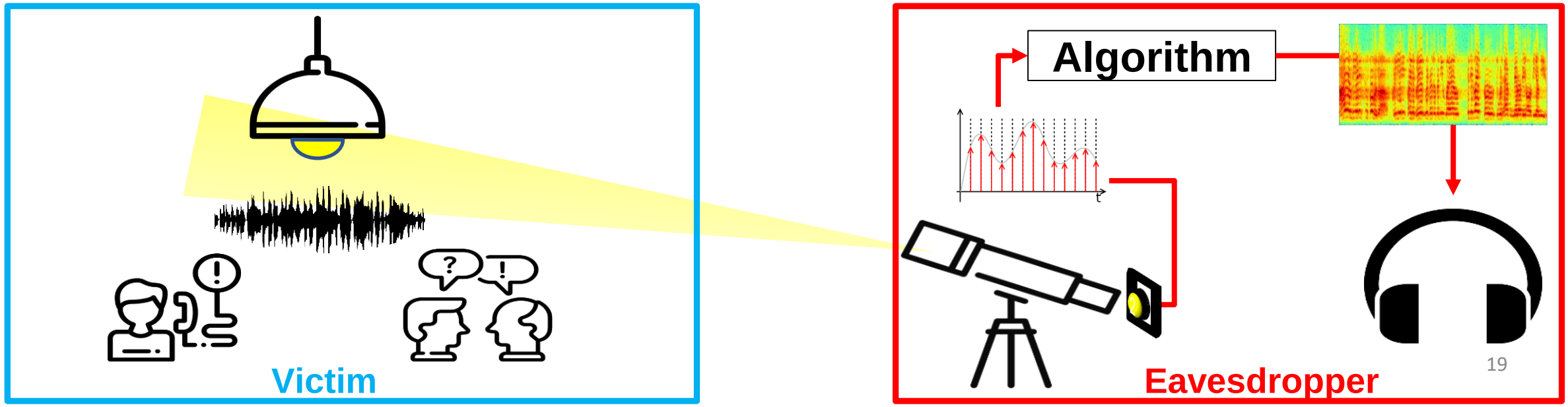
- It relies on an internal device
- It is active
- It cannot be applied in real time

BadVibes's Attack

BadVibe's eavesdropping attack is:

- **External**
- **Passive**
- Can be applied in **real time**

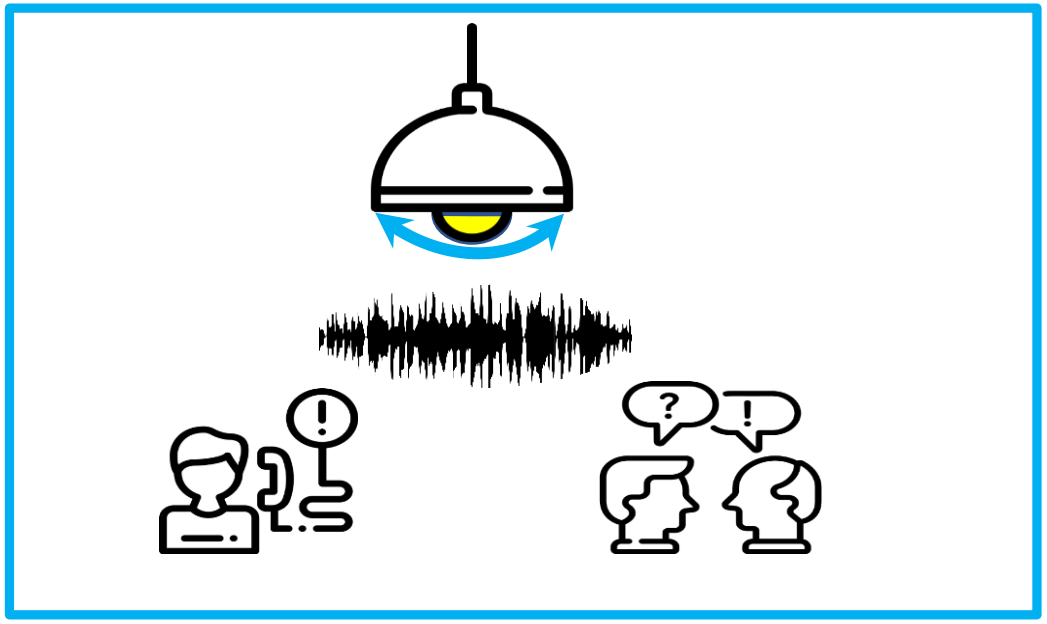
Overcoming the limitations of existing methods



Threat Model

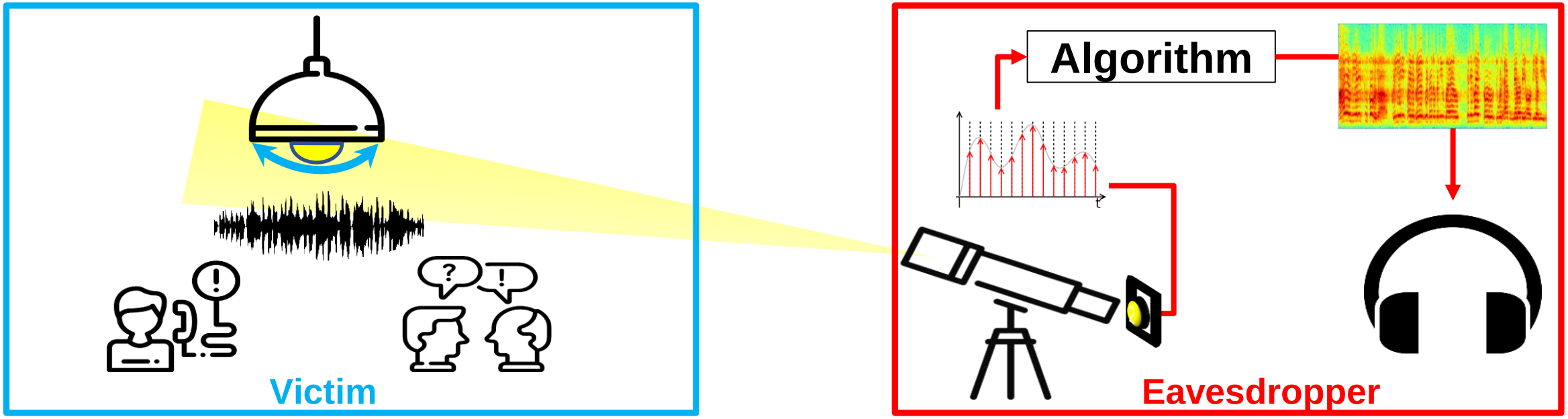
BadVibes's Threat Model

- We assume that a hanging light bulb exists in a target room.
- Sound waves (which are created as a result of a conversation) hit the surface of the light bulb.
- The light bulb is used as a diaphragm, converting sound into vibrations and as a transducer, coverts vibration to the light changes.



BadVibes's Threat Model

- An electro-optical sensor is directed at the light bulb using a telescope.
- The sensor is used as a transducer, converting the light bulb's vibrations to electrical current.
- The data from the sensor is sampled via an ADC and processed into an acoustic signal.



Physical Analysis

Physical Analysis

Light bulb vibrations caused by sound are not visible by the human eye.

Physical Analysis

Light bulb vibrations caused by sound are not visible by the human eye.

To get a better understanding of the light bulb's behavior, a few experiments were conducted:

- Measuring the vibrations
- Measuring light intensity as a function of the angle of the bulb
- Measuring light intensity as a function of the distance from the bulb
- Various bulbs types: Incandescent, LED, Fluorescent

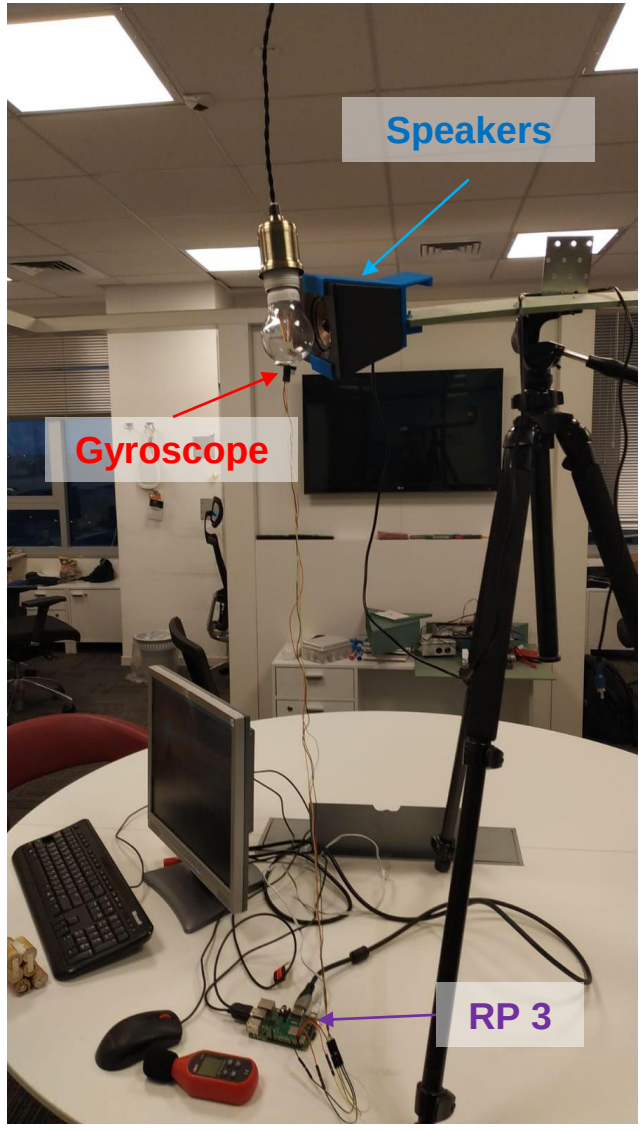
Measuring a Light Bulb's Vibrations

Experiment

- A **gyroscope** was attached to the bottom of a hanging light bulb (E27, 12 watts).

- We produced various sine waves in range of 100-400 Hz, at different volumes:
  - 70 dB – normal speech
  - 95 dB – loud conversation
  - 115 dB – air horn

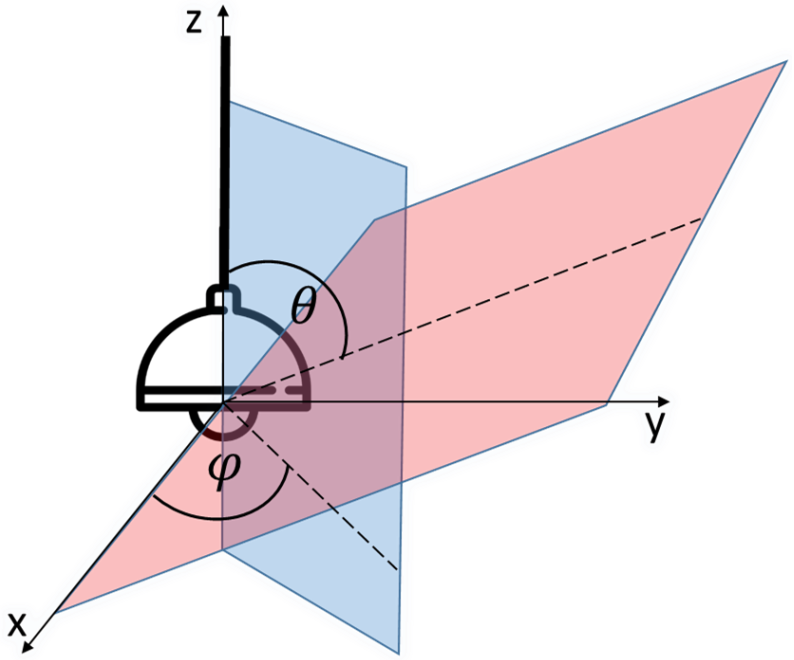
- We sampled the gyroscope at 800 Hz (using **Rpi 3**).

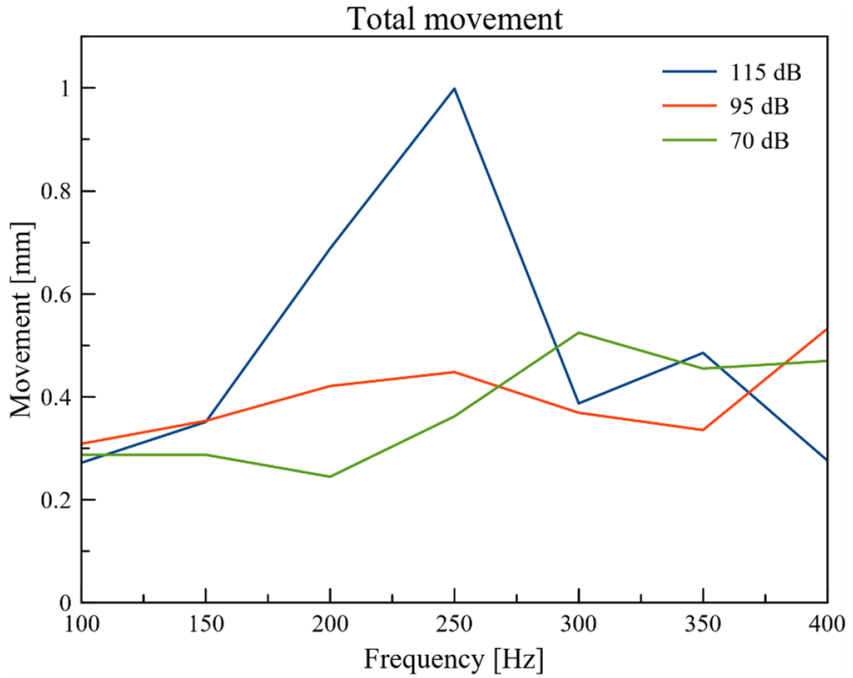


Measuring a Light Bulb's Vibrations

Results

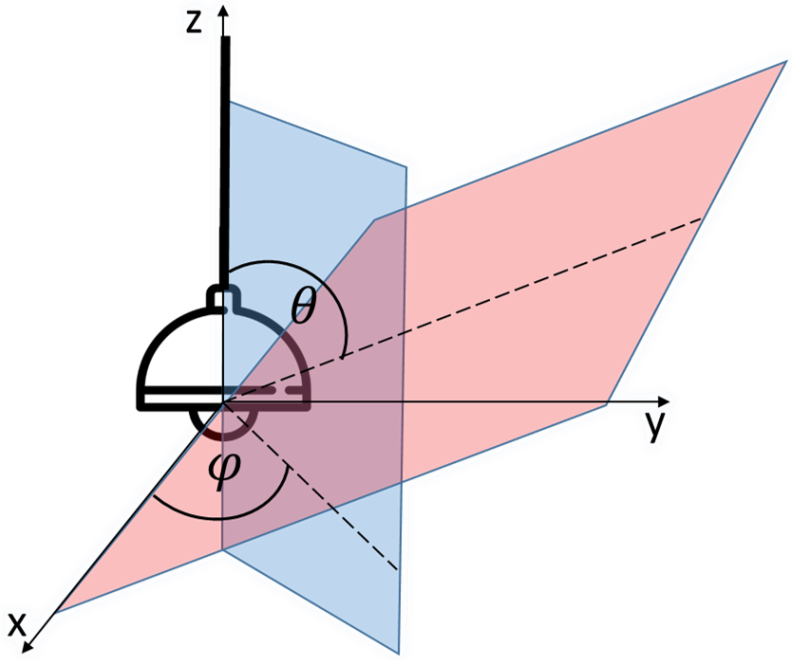
From the gyroscope measurements, we computed the total displacement of the bulb at the different frequencies.

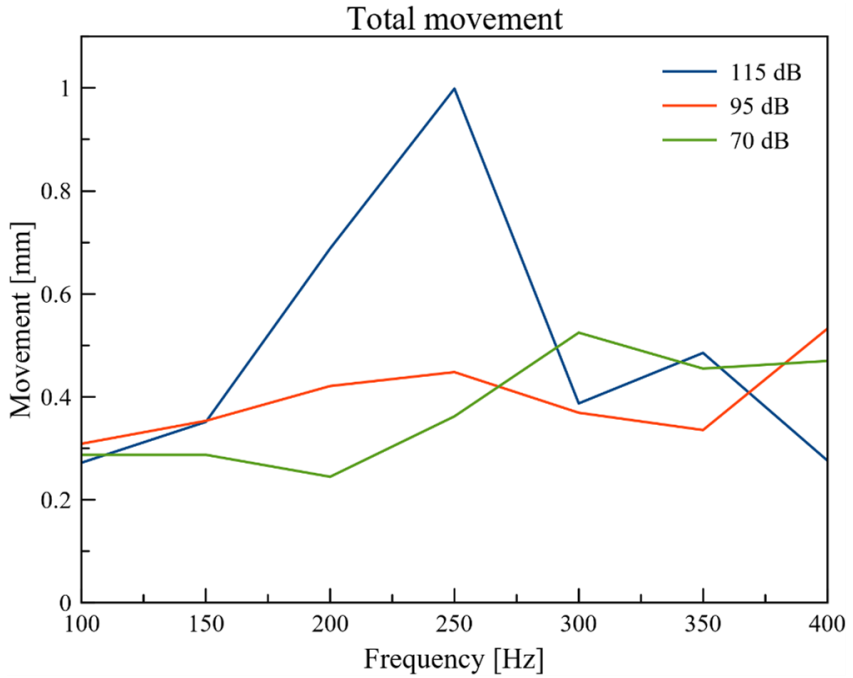




Measuring a Light Bulb's Vibrations

- The vibrations are between 300 – 1000 microns.
- The frequency response is not flat – the bulb responds differently at different frequencies.
- The vibrations are stronger as the sound volume increases.

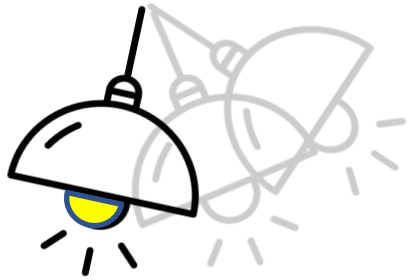




Light Intensity vs. Bulb Angle

Experiment

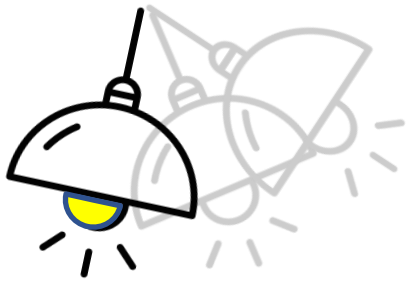
We directed an electro-optical sensor at an illuminated light bulb, and obtained voltage measurements from the sensor.

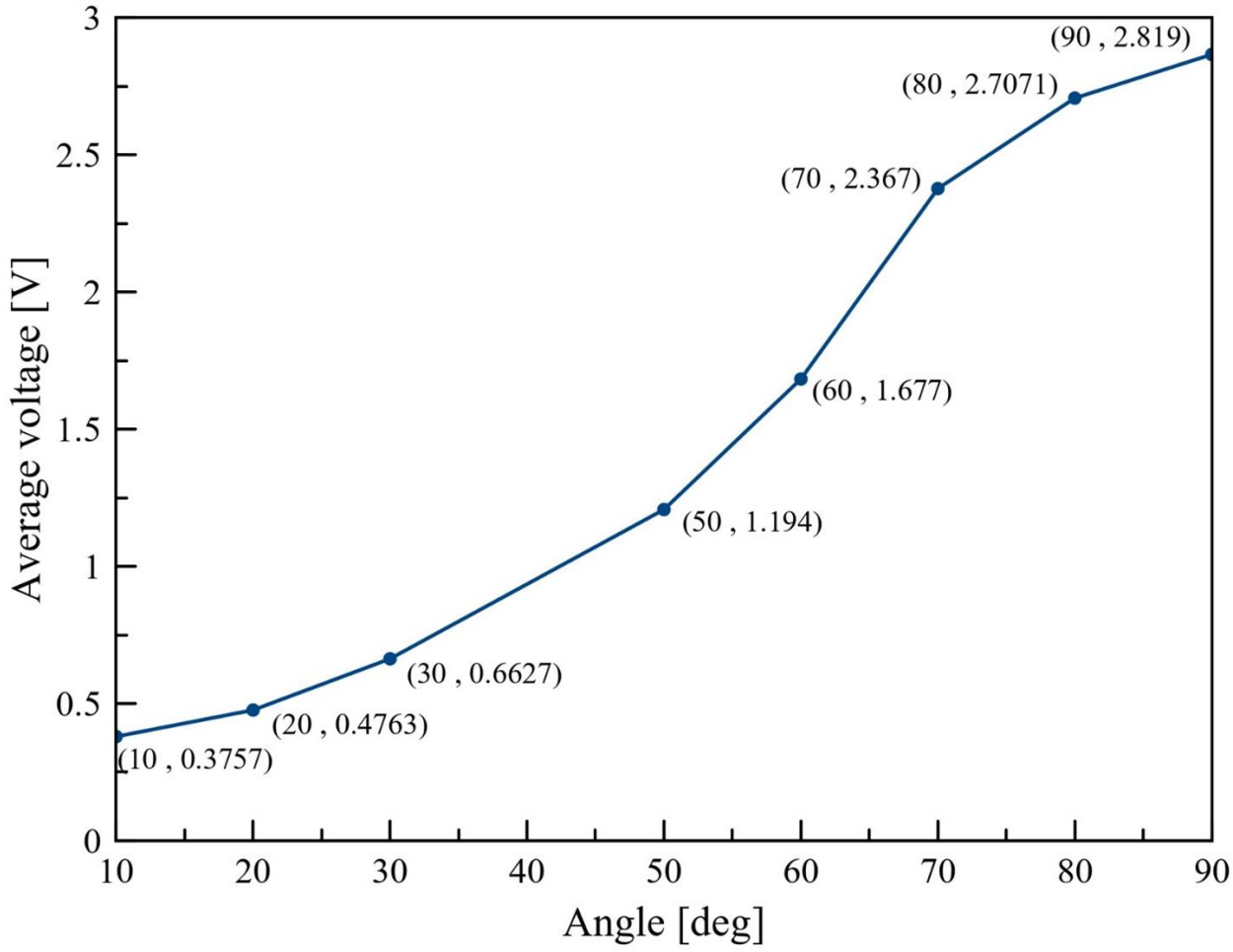


Light Intensity vs. Bulb Angle

Results

The light bulb is directional; as a result, the amount of light seen from different angles varies.

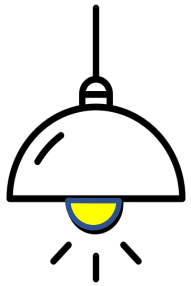




Intensity of Light vs. Distance

Experiment

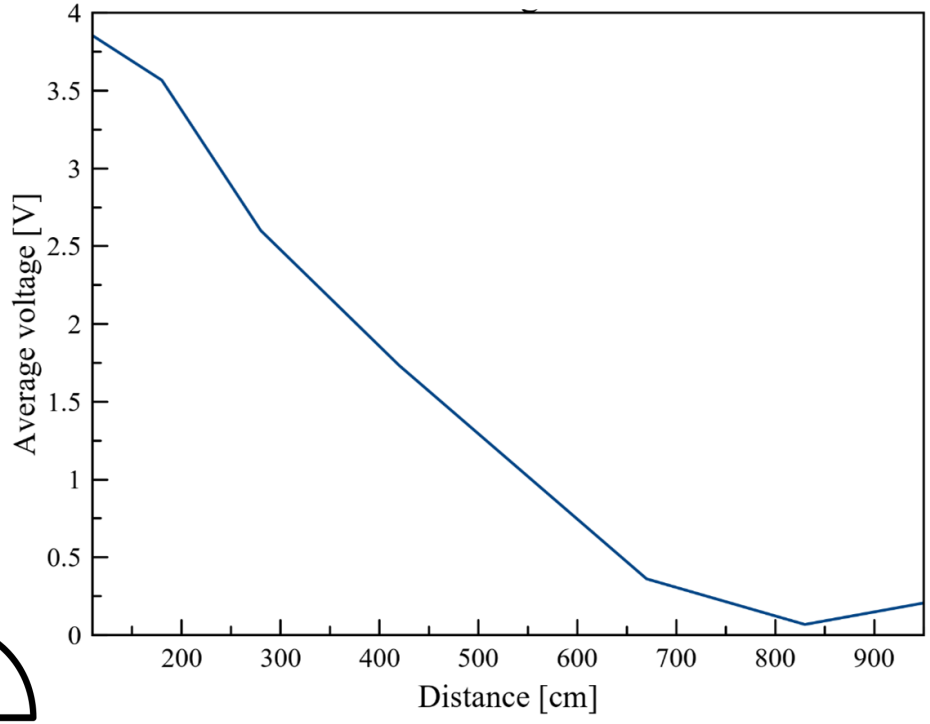
- We directed an electro-optical sensor at a hanging light bulb (when illuminated).
- We measured the voltage that was produced by the electro-optical sensor from various distances (100-950 cm).

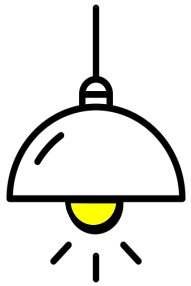


Intensity of Light vs. Distance

Results

A different amount of voltage is produced by the electro-optical sensor at different distances from the light bulb.





Setting the Criteria

Distance	Linear Equation	Expected Voltage Difference	
		at 0.3 mm	at 1 mm
200-300	y = -0.01x + 5.367	0.0003	0.001
300-420	y = -0.0062x + 4.3371	0.000186	0.00062
420-670	y = -0.0055x + 4.037	0.000165	0.00055
670-830	y = -0.0018x + 1.59	0.000054	0.00018

- A 16-bit ADC provides a sensitivity of 300 microvolts: $\frac{20}{2^{16}-1} \approx 300 \text{ microvolts}$

Setting the Criteria

Distance	Linear Equation	Expected Voltage Difference	
		at 0.3 mm	at 1 mm
200-300	y = -0.01x + 5.367	0.0003	0.001
300-420	y = -0.0062x + 4.3371	0.000186	0.00062
420-670	y = -0.0055x + 4.037	0.000165	0.00055
670-830	y = -0.0018x + 1.59	0.000054	0.00018

- A 16-bit ADC provides a sensitivity of 300 microvolts.
- A sensitivity of 300 microvolts is sufficient to recover the entire spectrum (100-400 Hz) from a distance of 200-300 cm.

Setting the Criteria

Distance	Linear Equation	Expected Voltage Difference	
		at 0.3 mm	at 1 mm
200-300	y = -0.01x + 5.367	0.0003	0.001
300-420	y = -0.0062x + 4.3371	0.000186	0.00062
420-670	y = -0.0055x + 4.037	0.000165	0.00055
670-830	y = -0.0018x + 1.59	0.000054	0.00018

- A 16-bit ADC provides a sensitivity of 300 microvolts.
- A sensitivity of 300 microvolts is sufficient to recover the entire spectrum (100-400 Hz) from a distance of 200-300 cm.
- In order to detect a bulb's vibration from 300 cm away: 1) the sensitivity of the system needs to be increased, or 2) the signal obtained needs to be amplified.

Comparison of Various Types of Bulbs

Experiment

We compared the SNR that was obtained from three types of E27 light bulbs:

Incandescent

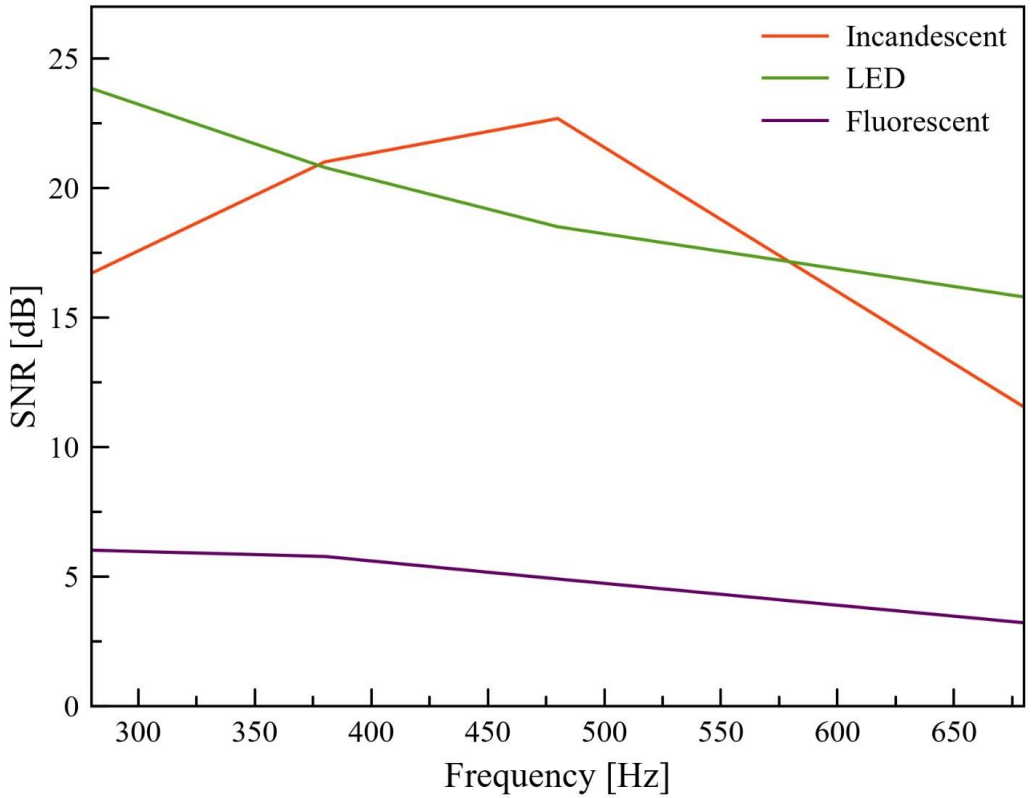
LED

Fluorescent

Comparison of Various Types of Bulbs

Results

1) Sound could be reconstructed from every type of hanging light bulb that was examined.
2) The SNR of incandescent and LED light bulbs is higher than the SNR of fluorescent bulbs.

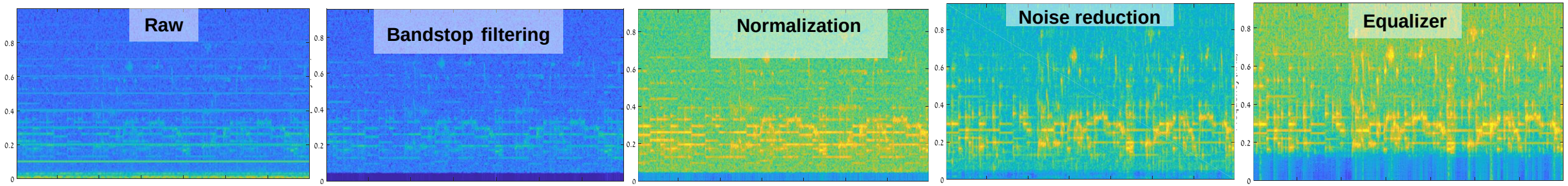


Algorithm

Algorithm

To reconstruct sound from the sampled optical signal, the digital data must be filtered from unwanted interference. In order to do so, the sampled data was processed in the following stages:

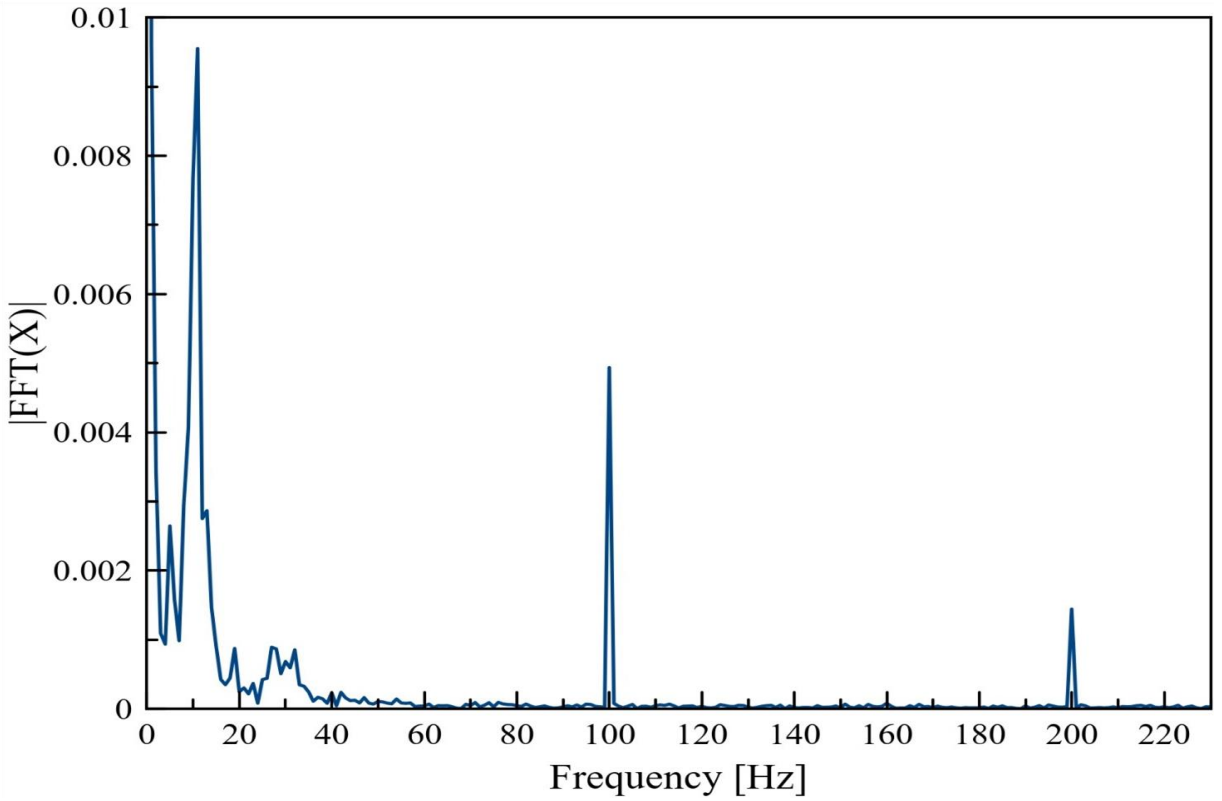
- Filtering using band-stop and high-pass filters
- Normalization
- Spectral Subtraction
- Equalization



Characterizing the Baseline: Band-Stop Filtering

Experiment

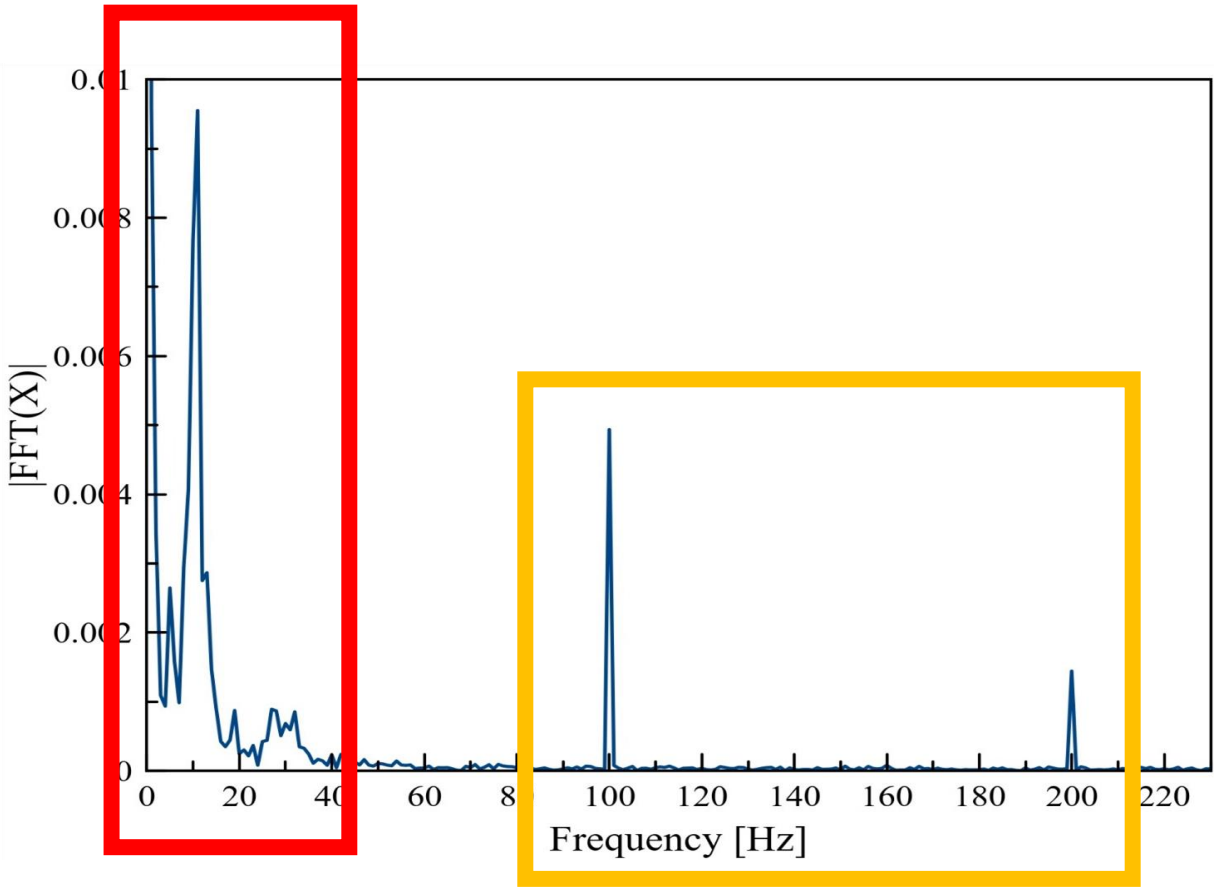
We obtained optical measurements via the electro-optical sensor when no sound is played near the light bulb.



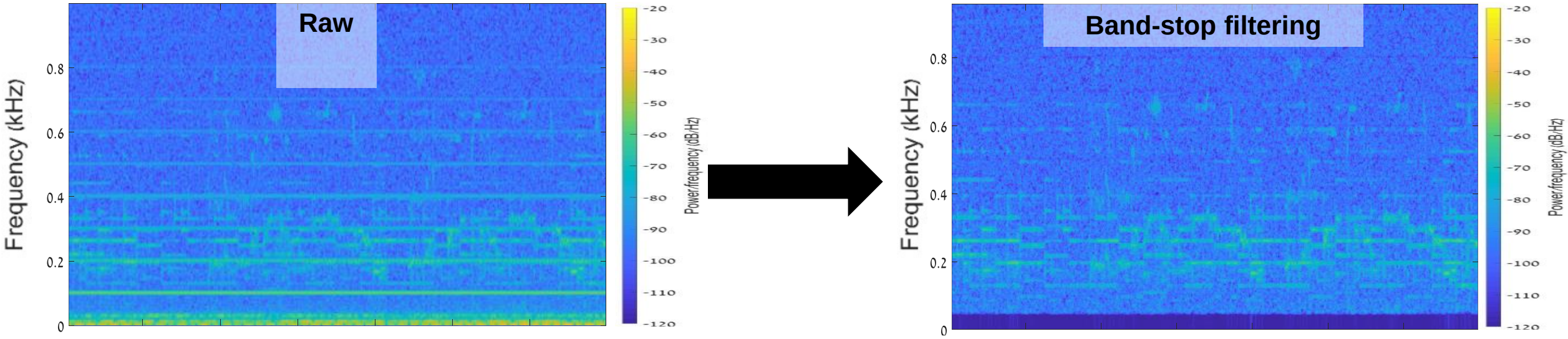
Characterizing the Baseline: Band-Stop Filtering

Results

- The LED bulb works at 100 Hz. There are peaks on the FFT graph at each of the harmonics (200 Hz, 300 Hz, etc.).
  - We need to filter this noise with band-stop filters.
- There is noise at low frequencies (below 50 Hz).
  - We need to filter this noise with a high-pass filter.



Characterizing the Baseline: Band-Stop Filtering

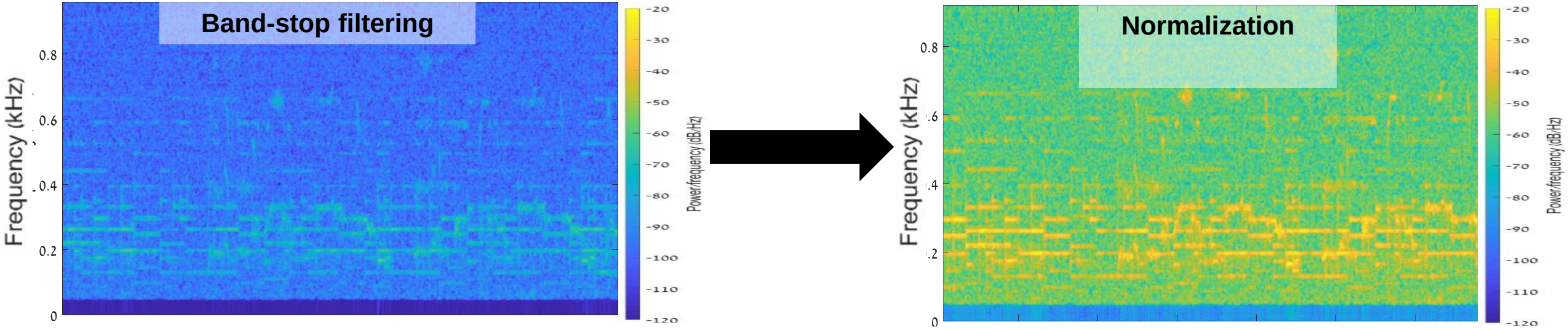


Normalization

To ensure the optimal use of the dynamic range available, the signal is **normalized**, by applying a constant amount of gain to bring the amplitude to a target level.

Normalization

To ensure the optimal use of the dynamic range available, the signal is **normalized**, by applying a constant amount of gain to bring the amplitude to a target level.



Spectral Subtraction

The ADC and electro-optical sensor are not ideal:

- Applying internal gain from the optical-sensor increases the spectral noise.
- The ADC has its own thermal white noise.

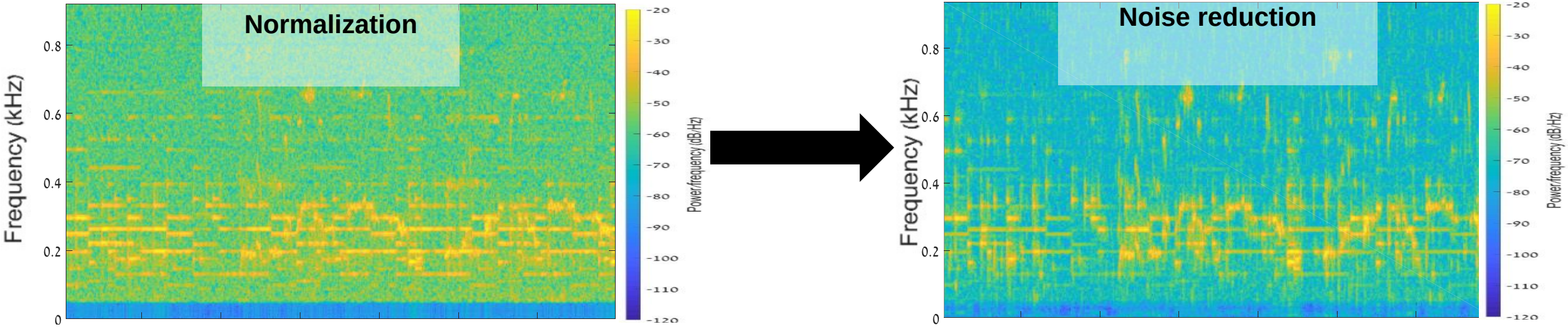
This results in a noisy audio signal.

Solution:

To help reduce the effects of unwanted noise, **spectral subtraction** is used.

In spectral subtraction, the noise power spectrum is estimated and subtracted from the signal.

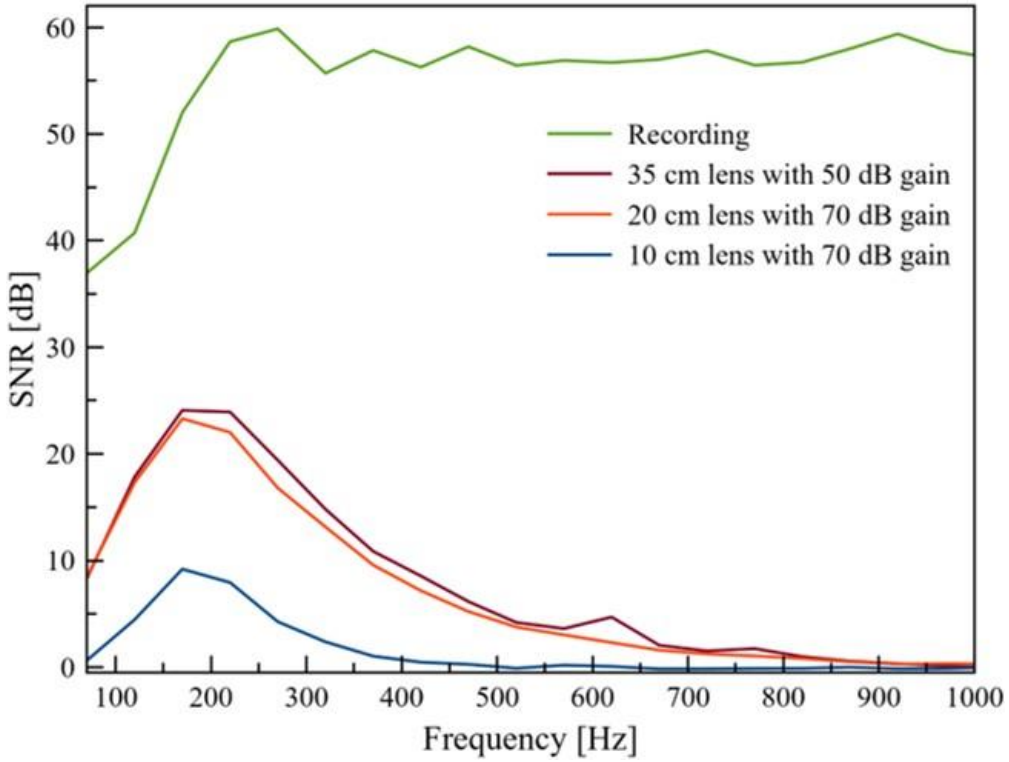
Spectral Subtraction – Results



Equalization

The power of the resulting signal is not equal across the spectrum.

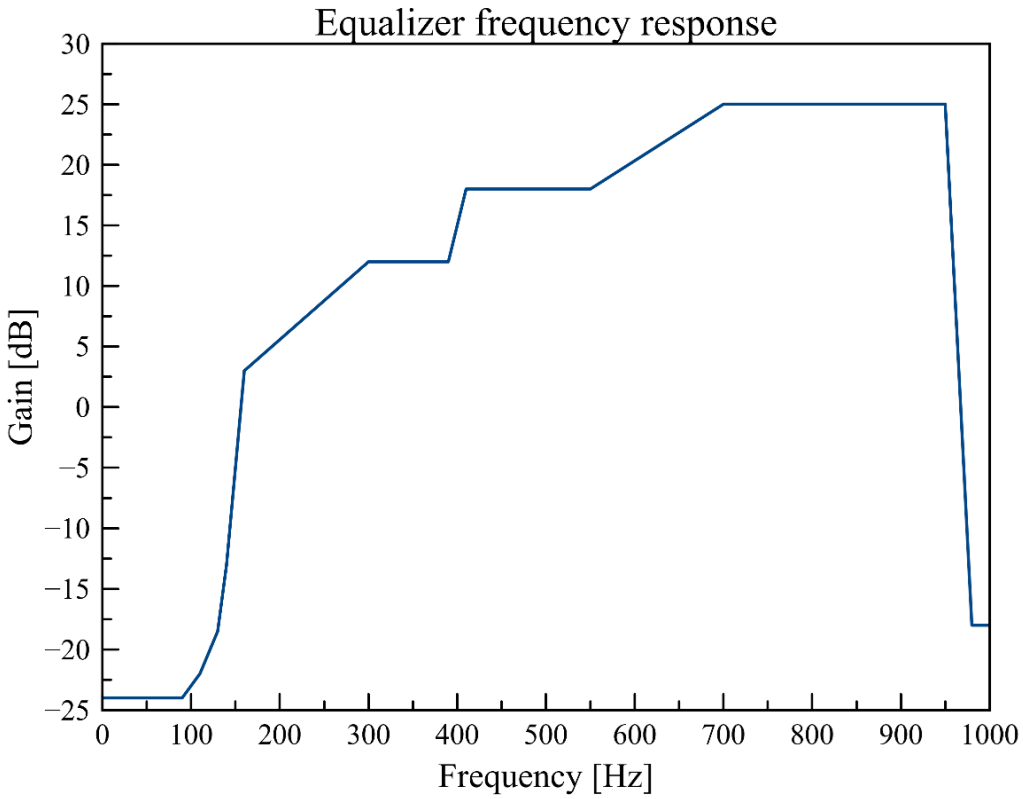
The experiments conducted showed that the SNR decreases at frequencies above 400 Hz.



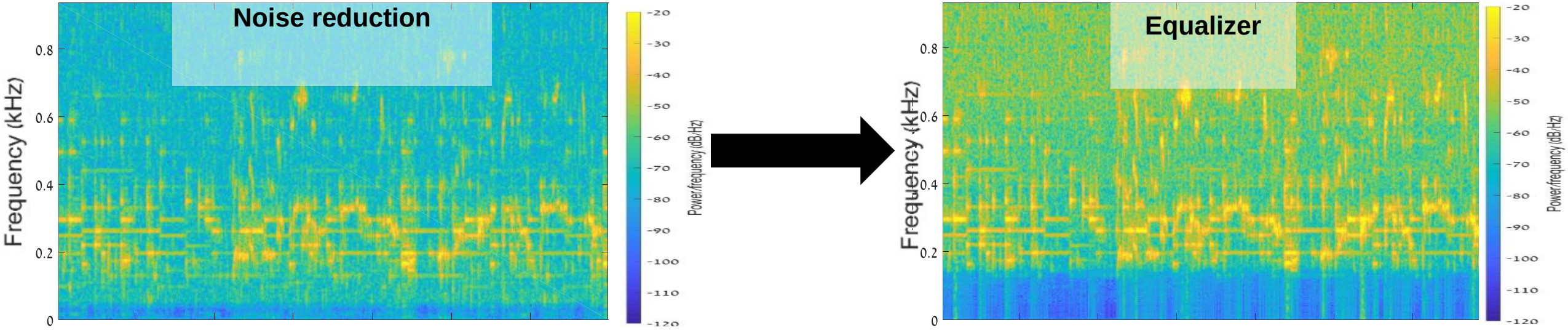
Equalization

Solution:

In order to compensate for power lost at high frequencies, an **equalizer** is applied.



Equalization – Results

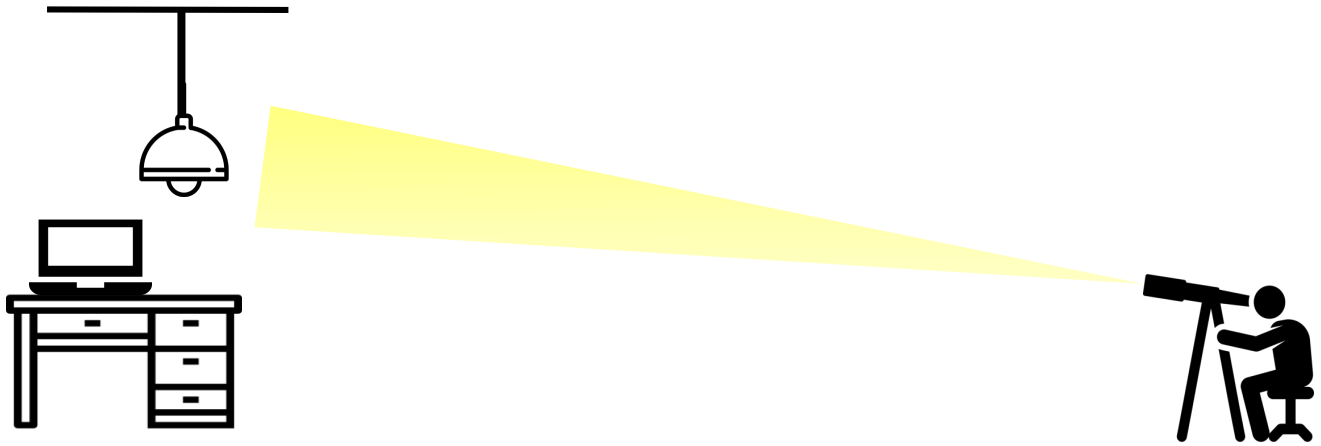


Evaluation

Experimental Setup

- In order to evaluate our attack's performance , we conducted two experiments:

1) Hanging bulb

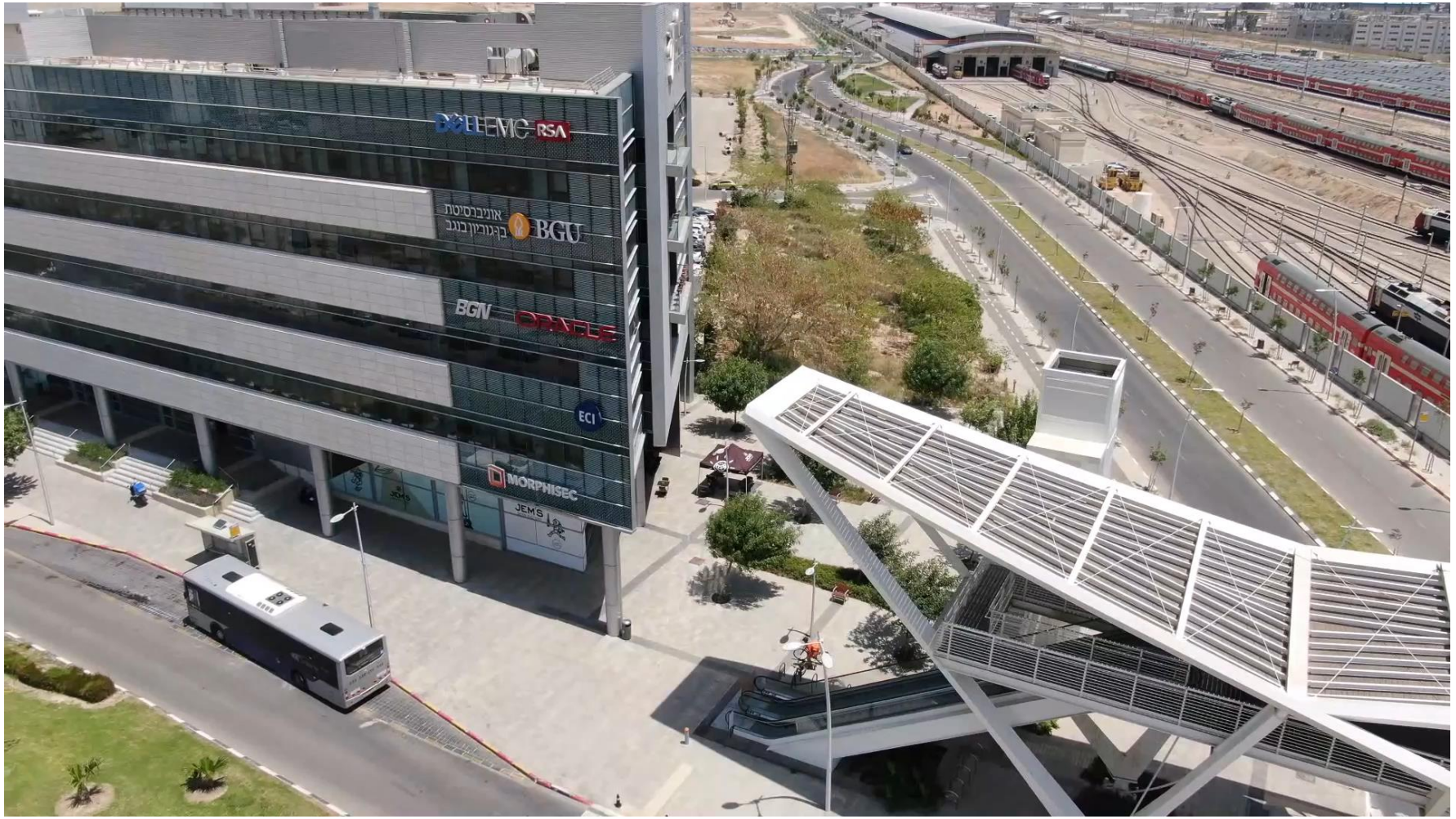


2) Table lamp

Experimental Setup

- In both of the experiments we tested our ability to reconstruct sound from 25m away from a target office.

Experimental Setup



The sound that was played inside the office cannot be heard from the bridge.

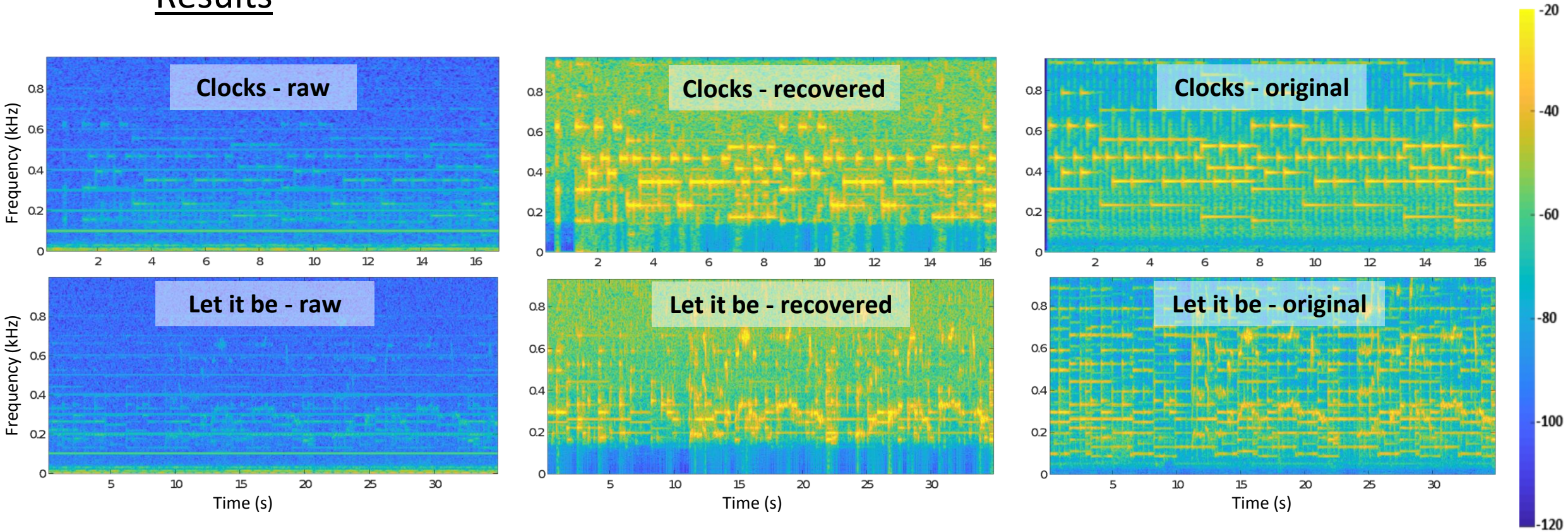
Evaluating the Eavesdropping Attack

First Experiment (optimized setup):

- We placed a **hanging light bulb** in the office with a direct line of sight through the window.
- The speakers were set to output sound at a volume of **110 dB**.
- Two famous songs were played inside the office:
  - "Let it Be” (The Beatles)
  - "Clocks” (Coldplay)
- A famous statement made by Donald Trump was played inside the office:
  - “We will make America great again!”
- The optical measurements were obtained.
- The signals were recovered.

Evaluation – Optimized Setup

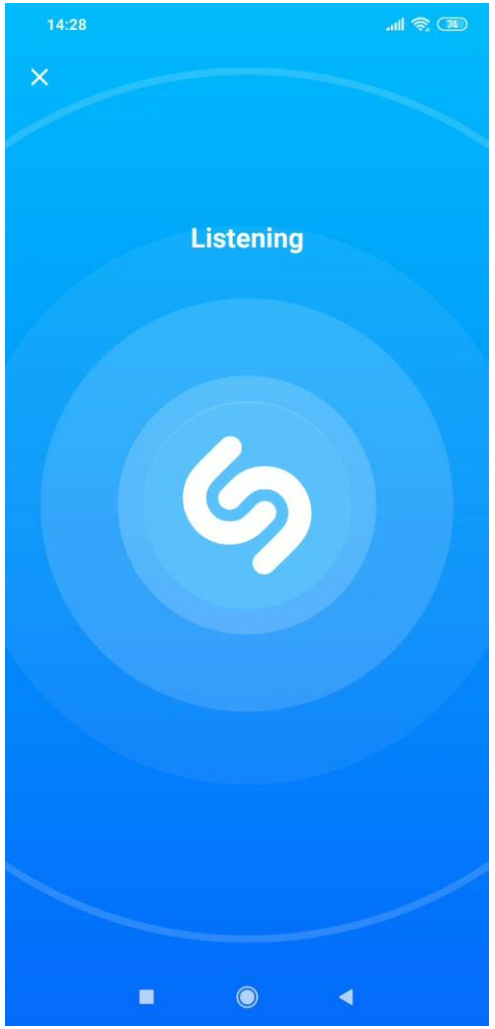
Results

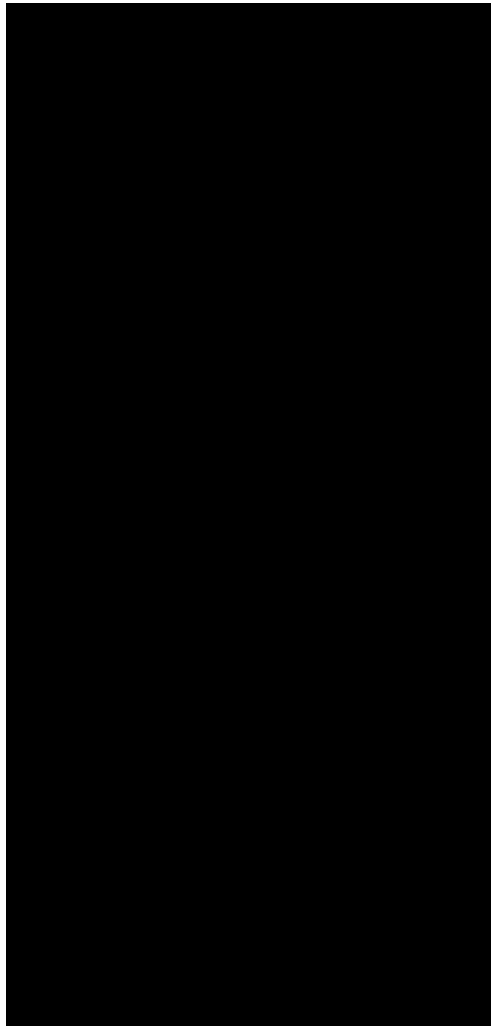


Evaluation – Optimized Setup

Results

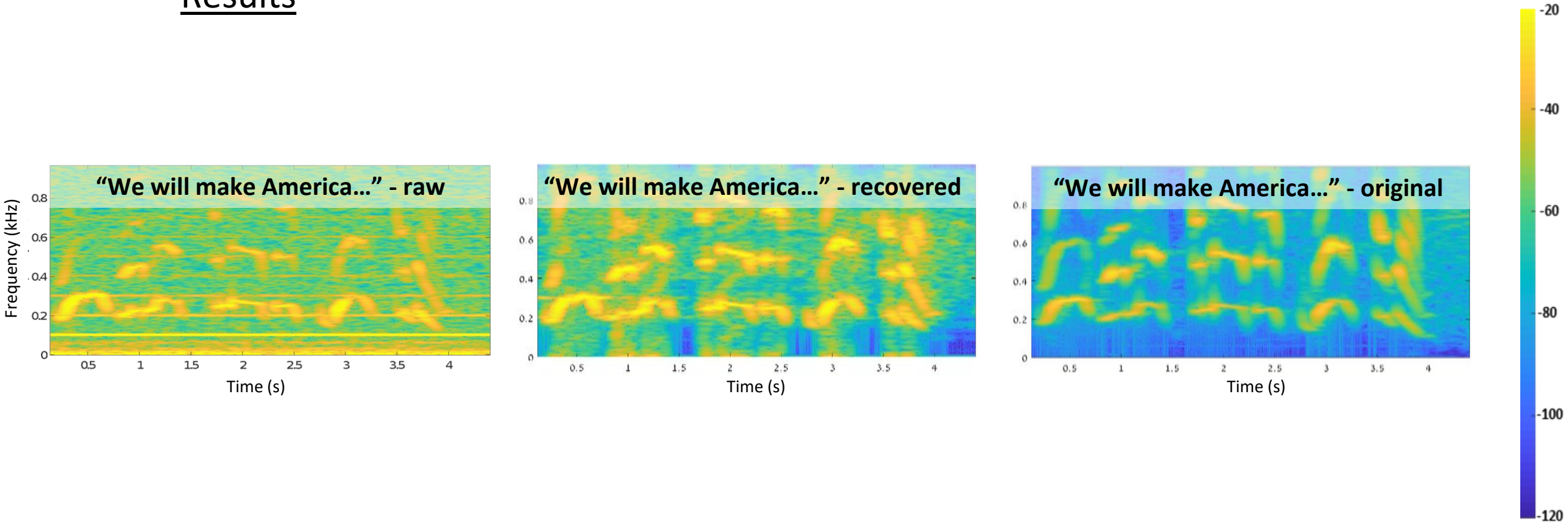
We Shazamed the recovered signals.





Evaluation – Optimized Setup

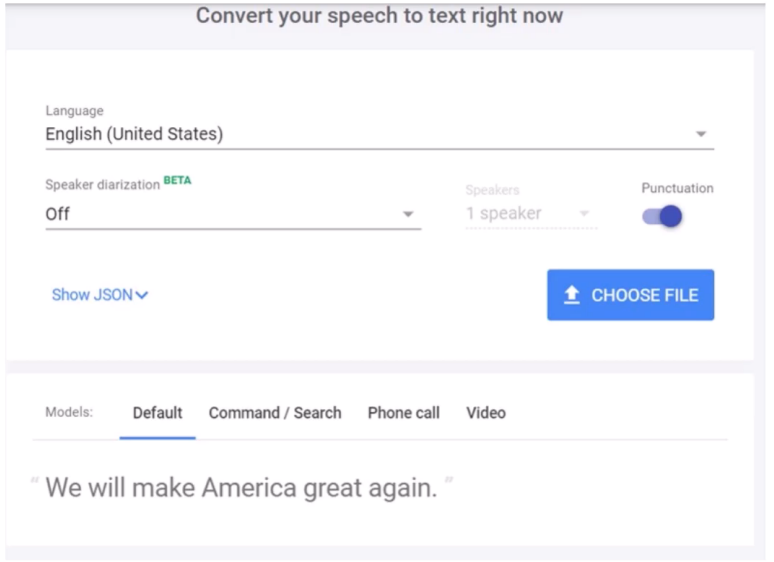
Results



Evaluation – Optimized Setup

Results

We investigated whether the recovered signal could be transcribed by Google's Speech-to-Text engine.



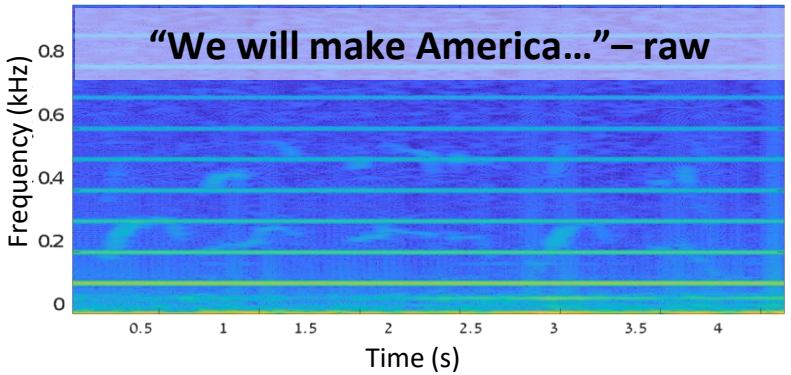
Testing the Attack

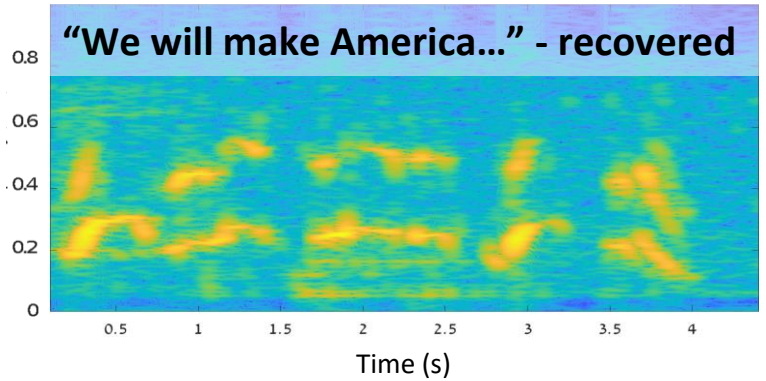
Second experiment (real setup):

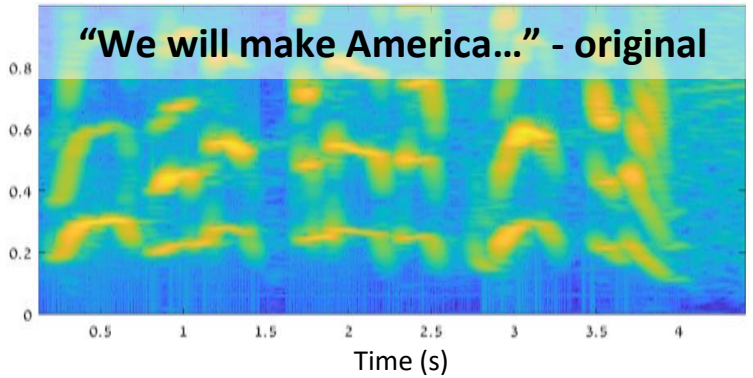
- We placed a **table lamp** in the office with a direct line of sight through the window.
- The speakers were set to output sound at a volume of **80 dB,** which is equal to a loud conversation.
- Donald Trump's famous statement was played inside the office:
  - "We will make America great again!"
- The optical measurements were obtained.
- The signals were recovered.

Evaluation – Real Setup

Results







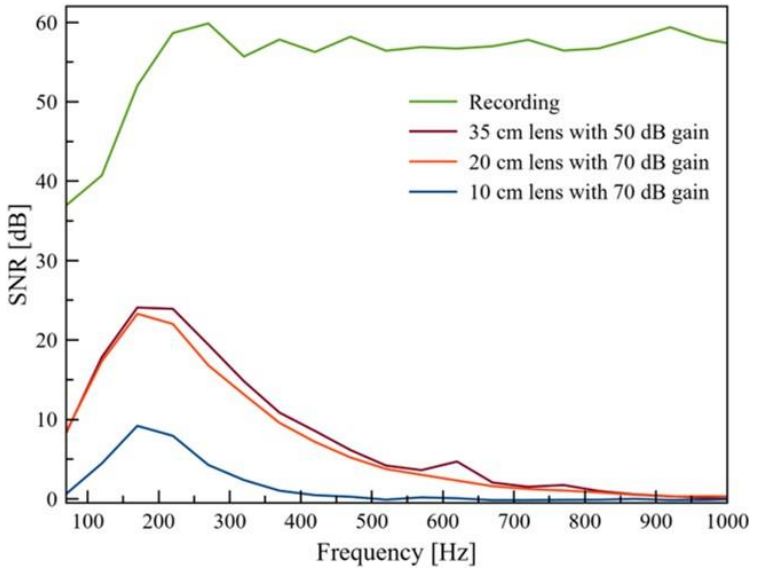
Potential Improvements

Potential Improvements

1. Telescope
   - Using a telescope with a larger lens diameter (*r*).

     <u>Why?</u>

     To increase the amount of light that is captured by the telescope.



Potential Improvements

1. Telescope
   - Using a telescope with a larger lens diameter (*r*).
2. Electro-Optical Sensor
   - Using a better (more sensitive) electro-optical sensor than the one we used (PDA100A2).
   - Using multiple electro-optical sensors directed at multiple light bulbs for multi-channel audio recovery.

Potential Improvements

1. Telescope
   - Using a telescope with a larger lens diameter (*r*).
2. Electro-Optical Sensor
   - Using a better (more sensitive) electro-optical sensor than the one we used (PDA100A2).
   - Using multiple electro-optical sensors directed at multiple light bulbs for multi-channel audio recovery.
3. ADC
   - Using an ADC with a lower noise level.
   - Using a 24/32-bit ADC instead of a 16-bit ADC.

Potential Improvements

1. Telescope
   - Using a telescope with a larger lens diameter (*r*).
2. Electro-Optical Sensor
   - Using a better (more sensitive) electro-optical sensor than the one we used (PDA100A2).
   - Using multiple electro-optical sensors directed at multiple light bulbs for multi-channel audio recovery.
3. ADC
   - Using an ADC with a lower noise level.
   - Using a 24/32-bit ADC instead of a 16-bit ADC.
4. Recovery Algorithm
   - Using advanced filtering techniques to filter noise (e.g., deep learning).

Thank You!

Questions?

Bibliography

[1] Y. Michalevsky, D. Boneh, and G. Nakibly, “Gyrophone: Recognizing speech from gyroscope signals,” in 23rd USENIX Security Symposium (USENIX Security 14). San Diego, CA: USENIX Association, 2014, pp. 1053–1067.

[2] L. Zhang, P. H. Pathak, M. Wu, Y. Zhao, and P. Mohapatra, “Accelword: Energy efficient hotword detection through accelerometer,” in Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services. ACM, 2015, pp. 301–315.

[3] S. A. Anand and N. Saxena, “Speechless: Analyzing the threat to speech privacy from smartphone motion sensors,” in 2018 IEEE Symposium on Security and Privacy (SP), vol. 00, pp. 116–133. [Online].

[4] Z. Ba, T. Zheng, X. Zhang, Z. Qin, B. Li, X. Liu, and K. Ren, “Learning-based practical smartphone eavesdropping with built-in accelerometer.” Proceedings of the Network and Distributed Systems Security (NDSS) Symposium 2020.

[5] M. Guri, Y. Solewicz, A. Daidakulov, and Y. Elovici, “Speake(a)r: Turn speakers to microphones for fun and profit,” in 11th USENIX Workshop on Offensive Technologies (WOOT 17). Vancouver, BC: USENIX Association, 2017. [Online]. 14.

Bibliography

[6] N. Roy and R. Roy Choudhury, “Listening through a vibration motor,” in Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, ser. MobiSys ‘16. New York, NY, USA: ACM, 2016, pp. 57–69.

[7] A. Kwong, W. Xu, and K. Fu, “Hard drive of hearing: Disks that eavesdrop with a synthesized microphone,” in 2019 2019 IEEE Symposium on Security and Privacy (SP). Los Alamitos, CA, USA: IEEE Computer Society, May 2019.

[8] R. P. Muscatell, “Laser microphone,” Oct. 25 1983, US Patent 4,412,105.

[9] Davis, A., Rubinstein, M., Wadhwa, N., Mysore, G. J., Durand, F., & Freeman, W. T. (2014). The visual microphone: Passive recovery of sound from video.