VOICE DOSIMETER

By

David Gong (dsgong3)

Jaden Li (sizhel2)

Michael Rizk (rizk2)

Final Report for ECE 445, Senior Design, Spring 2025

TA: Chi Zhang

07 May 2025

Project No. 79

Abstract

Our group has created a voice dosimeter that is capable of recording and measuring the user's sound pressure level, fundamental frequency, and cepstral peak prominence to within 10% accuracy. With the choice of a vibration sensor, our device measures only the user's vocal cords, ensuring total isolation from external sources of sound. Furthermore, using a 3V button battery and low-power components, our group's voice dosimeter is capable of continuously transmitting data via Bluetooth Low Energy for more than 8 hours.

Table of Contents

I. Introduction

Problem

Professional voice users – teachers, singers, actors, actresses, broadcasters, clergy, salespeople, courtroom attorneys, telemarketers, and health care specialists – constitute about 30% of the working population [1]. Many of these people will commonly experience voice disorders associated with high voice usage that can negatively affect productivity. The US societal costs of voice-related teacher absenteeism and treatment expenses alone have been estimated to be as high as 2.5 billion dollars annually [2]. Such absenteeism could be prevented for teachers and other voice professionals if they were able to measure how much they were using their voices every day through the use of a voice dosimeter. Measurement of voice usage can also be used in the recovery of voice disorders, as vocal rest is often recommended. Despite the apparent utility of voice dosimeters, there are currently none that are commercially available. Some devices were available in the past, but they cost thousands of dollars. A low-cost and widely available voice dosimeter would allow for clinical and research use of voice-related problems and voice therapy options.

This project is conducted in collaboration with graduate student Charlie Nudelman and Professor Pasquale Bottalico at the College of Applied Health Sciences. Their group had previously developed a DIY voice dosimeter using a contact microphone and a portable audio recorder. They are still using this device today, but there are a few improvements they would like to see.

Solution

The device that Charlie and Pasquale are using is bulky and requires a wired connection. It is impractical for patients to wear daily and collect data for a long period of time. We aim to create a cheaper and more comfortable voice dosimeter that is capable of recording data for long periods of time without recharging while the data can be uploaded to another device wirelessly.

To achieve this goal, we will be using a vibration sensor instead of a contact microphone that Charlie and Pasquale's initial design used. This will allow us to record the vibrations from the user's vocal cords with a smaller form factor. Afterward, this data can either be transmitted via Bluetooth onto an external computer. Finally, this data will be processed to extract critical information about vocal usage and strain, which researchers can then analyze to improve voice therapy options.

Visual Aid

As shown in Figure 1, the voice dosimeter will be a small device that can comfortably be attached to the vocal cord above the collarbone. It will record the user's vocal vibrations as they speak.

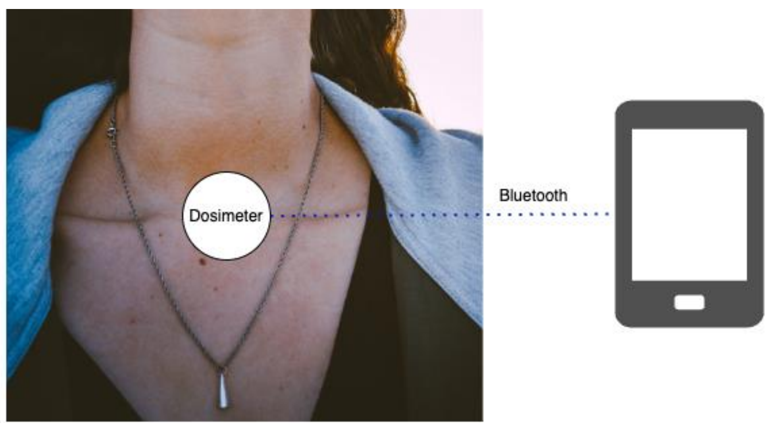

Figure 1: Visual aid of voice dosimeter system

High Level Design

Our device is controlled by a microcontroller from the STM32WB family of MCUs. Connected to the MCU is the vibration sensor that detects vibrations from the user's vocal cords, an RF system to transmit data wirelessly, and a 4-gigabit NAND flash memory to store data. All components can run on 1.8 V, which is supplied by our power management subsystem that steps down a 3.0 V button battery to 1.8 V.

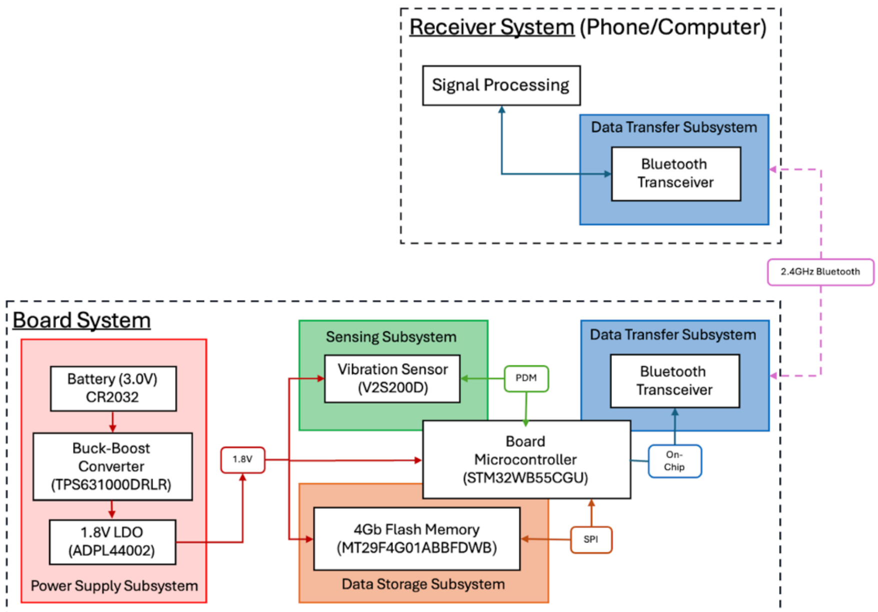

Figure 2: High-level block diagram of voice dosimeter system

High-level Requirements

1. The device shall accurately measure voice features including sound pressure level within 2dB, fundamental frequency within 5Hz, and cepstral peak prominence within 2dB.
2. The receiver phone or computer shall have a website for viewing voice features in real time with less than 1 second of latency.
3. The device shall have 8 hours of battery life to last through a full workday.

II. Design

Electrical Subsystems Overview

Power Management

One of the most critical subsystems of the device is power management. As this is a wearable device, we want to ensure that even with a small battery, our device is able to run continuously for at least 8 hours. Thus, we will use a 3.0 V button battery, stepped down to 1.8 V for the supply level, as seen in Figure 2.

To achieve the desired supply voltage, we use a buck-boost converter in series with an LDO. As an LDO steps down voltage through heat dissipation, we wanted to step down the voltage from 3.0 V more efficiently through a buck converter with a high switching efficiency. After converting the voltage to 1.9 V, an LDO is used to reduce switching noise. The buck-boost converter can also step up the battery voltage to the LDO input if the battery is depleted below 1.9 V.

Vibration Sensing

The sensing subsystem is the primary driver of this device. We use the V2S200D vibration sensor from Syntiant to measure the vibration of the user's throat as they speak. By using a vibration sensor instead of a microphone, external sounds and vocalizations are significantly reduced. The data collected from the accelerometer will be sent to the microphone via 16-bit pulse density modulation (PDM).

At first, our group chose the BMA580 accelerometer from Bosch. It was smaller and had a much lower operating current, in the microamp range. However, as this device had to be hand-soldered, the BMA580 was too small (1.2mm x 0.8mm) to reliably solder so that it functions properly. Thus, we switched our sensor to the V2S200D vibration sensor, which has a larger form factor (3.30mm x 2.30mm).

MCU/Bluetooth

The microcontroller chosen for this device is the STM32WB55CGUx, which has an onboard Bluetooth transceiver. In this system, the Bluetooth low-energy stack on the MCU functions as the peripheral device, whereas the Bluetooth module on a phone or computer functions as the central device. When the data from the vibration sensor is stored in the data buffer, the MCU can either store the data in the device's memory IC or transmit the data via Bluetooth Low Energy (BLE). An external antenna and RF front-end consisting of a pi matching network and a low pass filter are also necessary.

It is possible to achieve BLE transmission with another MCU and an external Bluetooth module. However, ST's STM32WB family of microcontrollers was designed with wireless communication in mind. It consists of two cores: the M4 cortex and the M0+ cortex. The M4 cortex handles the device's custom application, whereas the M0+ cortex handles the BLE stack in the background.

Storage

The storage subsystem consists of a 4-gigabit flash memory IC. The flash memory has enough storage for at least eight hours of raw accelerometer data. Initially, we had planned to use the flash memory to store data when the dosimeter was not connected via BLE. However, it would take too long to transmit 8 hours of data with BLE. While we were able to verify the requirements of the flash memory, we did not include it in our final MCU firmware.

Circuit Schematic

The circuit schematics and PCB layout for our device is listed in Appendix C.

Software Overview

Firmware Flowchart

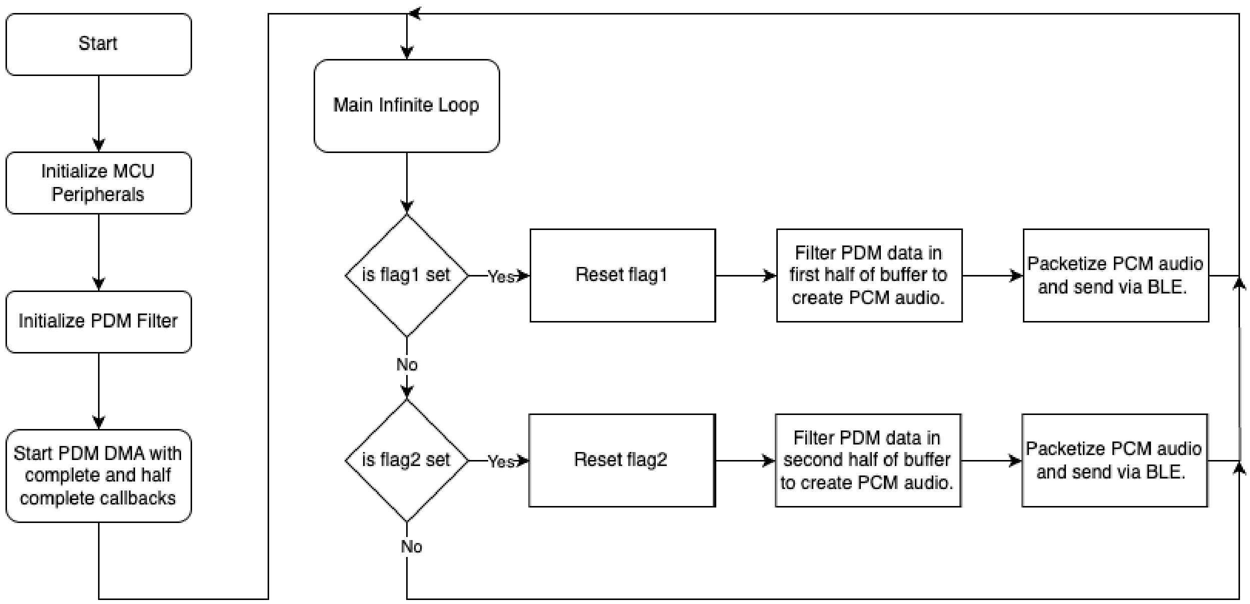

Figure 3: The flowchart for the microcontroller firmware

Firmware Description

The first step of the firmware after booting is to initialize the MCU peripherals. These include GPIO, DMA, UART, Serial Audio Interface (SAI), Quad SPI, ADC, and RF. Next, we enable the PDM filter [3]. We use a 2 MHz clock for our vibration sensor, and the PDM filter decimates the signal by a factor of 256, giving us an audio sample rate of 7812.5 Hz. Next, we start the PDM circular DMA and enable the callbacks. There are two callback functions for when the PDM buffer is half complete and fully complete that set flags 1 and 2, respectively. Finally, we enter our main loop where we filter the audio and transmit it via BLE. Using flags 1 and 2, we are able to process the first half of the buffer while the second half is being filled and vice versa. This enables us to never miss any data.

Website Overview

Figure 4: Voice Dosimeter web application

The website consists of a JavaScript React front-end and a Python backend. On the website, we have Bluetooth, SPL calibration, and live vocal feature plotting capabilities. The user can also download the audio for their own analysis. We chose to create a website instead of a desktop application because a website will work on both computers and cell phones.

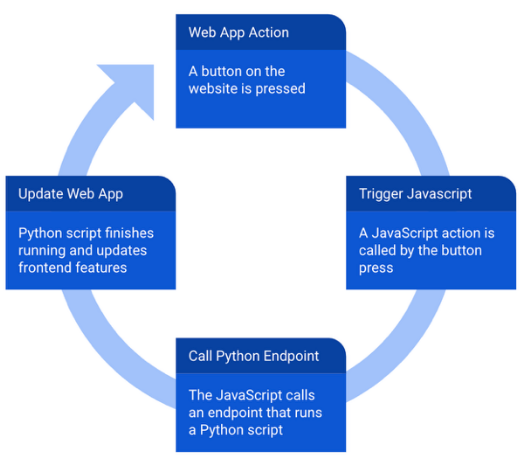

Figure 5: Website Flowchart

Calibration

The SPL will need to be calibrated each time the device is removed and reattached due to the differences in placement or contact affecting mechanical coupling. To calibrate the device, the user will place an SPL meter or SPL meter phone app 0.5 meters away from the mouth. Then, the user will make a long A vowel sound and just the SPL calibration setting on the website to match the SPL meter readings.

Processing

The processing subsystem will estimate the following features of the user's voice in real time: sound pressure level (SPL), fundamental frequency (f0), cepstral peak prominence (CPP), and the vocal doses. These features are extracted from the raw accelerometer data and will be displayed accordingly in a web application. SPL is calculated using total energy, which is dependent on the calibration process. F0 is estimated using the Praat autocorrelation algorithm. The CPP is calculated using the algorithm in Hillenbrand et al [4]. We will take the Fourier Transform of the accelerometer data to obtain the Power Spectrum. Then we will take the square of the log of the Power Spectrum and then take the Fourier Transform of that again. Finally, we take the difference in amplitude from a peak prominence to the baseline, and that defines the CPP.

Vocal doses are calculated using a weighted sum of the three other values. There are different types of vocal doses based on the weights. The Vocal doses that we will calculate based on the algorithms in Assad et. al. [7] are the following: Dynamic dose, Energy Dose, Radiated Dose, and their normalized forms. The Dynamic Dose indicates the total time of vocal fold vibration in the workday. The Energy Dose is the quantitative descriptor of energy delivered to the vocal folds. Radiated Dose represents the sound energy emitted from the vocal folds during vocal activity. The normalized forms are just normalized by the total duration of vocal activity (this is needed because during an eight-hour window, no one ever speaks for the entire time).

Mechanical Overview

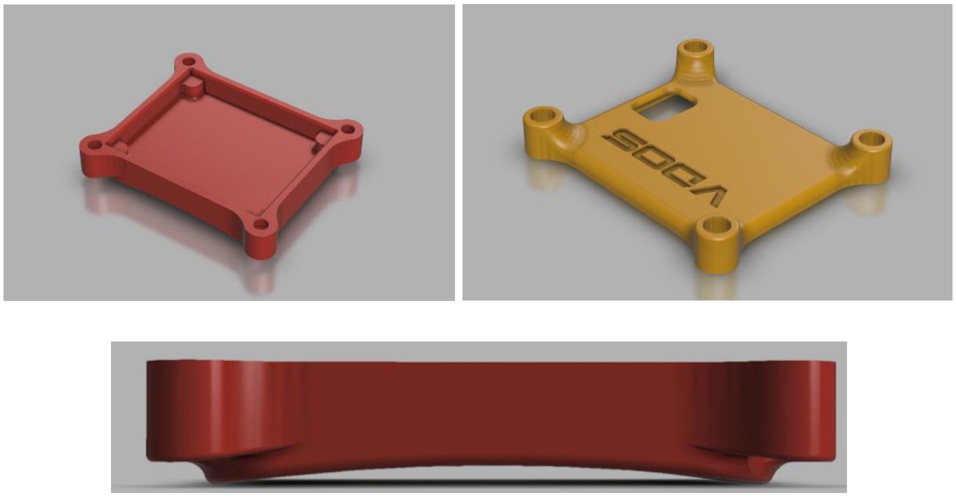

Figure 6: Diagram of mechanical design

PCB Enclosure

To house the PCB, we used a custom-designed, 3D-printed enclosure made from TPU filament with a Shore hardness of 95A. The slight curve on the bottom was made in order to comfortably fit the average radius of curvature for an adult male’s neck. However, it is highly likely that this housing would be comfortable for most adults. We chose 3D-printing over a silicone exposure because of the ease and reproducibility of manufacturing.

Medical Dressing Tape

The enclosure is attached to the skin using a transparent dressing. The tape is able to stay on for more than 8 hours.

III. Design Verification

All our subsystem requirements and verifications are detailed in appendix A. We performed the verifications and satisfied all the requirements except for one that is detailed below.

The tests for our device were done in the partially anechoic chamber in the UIUC Speech and Hearing Sciences building. We tested our voice dosimeter in that room and compared it with the results from a measurement microphone. For controlled testing, the tester read a passage in various background noise settings. Most notably, we were interested in comparing the performance of the microphone and the voice dosimeter in a noise-free setting and in white background noise.

Most of the sensing requirements we had were verified completely. However, one of the requirements that we failed to achieve was the accuracy of our sensor, as seen in Figure 7

Clean Passage	SPL Average (dB)	F0 Average (Hz)	CPP Average (dB)
Microphone	77.5	103.8	12.9
Voice Dosimeter	86.1	106.6	11.9

Passage with white noise	SPL Average (dB)	F0 Average (Hz)	CPP Average (dB)
Microphone	80.7	108.1	12.9
Voice Dosimeter	85.7	100.5	11.9

Figure 7: Results from tests in the UIUC Speech and Hearing Science Building

We were not able to achieve our high-level requirements for accuracy for the SPL Average and the fundamental frequency. There are some reasons as to why we suspect that there was a discrepancy between the devices. The first of which is that both of the devices were sometimes prone to octave errors, which would've impacted the f0 values. Also, we know that the microphone was less noise resistant than the accelerometer-based voice dosimeter, which could've led to the closer F0 readings in clean passage verses the passage read with background white noise. Lastly, the accelerometer cannot pick up certain vocal sounds such as unvoiced consonants which are not produced in the vocal folds.

IV. Costs

Cost Analysis

The cost for electronic components for a single device is detailed in Appendix B, with a subtotal of $31.92. Material cost for development, detailed in Appendix B, was $259.71. Furthermore, the average salary of a UIUC electrical engineering graduate is around $90,000 a year, which amounts to about $40/hr. Each of the three group members contributed 100 hours to the project. An estimated labor cost of $40/hour x 2.5 x 100 hours gives a total of $10,000 per person. This brings the cost of labor to $30,000 for the entire team. Altogether, the total cost for the research and development of this product is $30,259.71.

V. Ethics and Safety

As we proceed with this project, we want to acknowledge the ethical and safety considerations that are relevant to the IEEE and ACM code of ethics, as well as the steps we have taken to mitigate the concerns.

To begin, our device has been designed with health research in mind. Our device will record the user’s vocal data, which can be considered personal health information. Thus, we want to ensure that such data is handled responsibly and only accessible to the relevant, permitted parties. When data is collected, it will either be temporarily stored on the device or transmitted to a computer, where it will be stored and analyzed. As the data on the device and external computer would be difficult to access without the user’s permission, the point when the data is most vulnerable is during data transmission. To prevent potential information leaks, our data will be transmitted via Bluetooth, which follows standard security protocols to ensure a secure connection.

Furthermore, as per ACM Code of Ethics Section 1.6, everyone is entitled to their privacy, which must be respected. As our device records vocal data, there is the consideration that the voices of surrounding individuals could be recorded. However, our device does not use a microphone to record the user’s voice; instead, we use an accelerometer placed directly on the user’s vocal cord. This ensures that only the user’s vocal data is recorded, preventing the potential invasion of privacy of other individuals. Since not all vocal sounds are picked up by the accelerometer, the speech is not intelligible. This ensures the privacy of the user as well.

Finally, there is also the concern of wearable electronics. Not only is our device designed to be directly attached to the user’s skin, but our device also utilizes a lithium-ion battery. We must consider the potential hazards and dangers that this device may bring. Adhering to IEEE Code 7.8.9, we deliberately chose a low-voltage battery of 3V. Furthermore, the internal circuitry has been designed to operate at 1.8V and low currents, which further mitigates the risk to the user. Additionally, our wearable device will be enclosed in silicone for comfort, doubling as an insulating layer for protection.

VI. Conclusion

Our Voice Dosimeter met most of our requirements, and it is well-suited for the task it was designed for. As the introduction mentions, vocal disorders can be very common among the working class, and there really isn't an available device on the market. With a comfortable and IEEE-compliant design and an easy-to-use UI, our voice dosimeter offers an affordable and secure platform for monitoring vocal health. We are confident that a device like ours can make a big impact in diagnosing vocal disorders and consequently ensure a healthier and more productive workforce.

References

[1] R. W. Chan, I. R. Titze, and M. R. Titze, ‘Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa’, *The Journal of the Acoustical Society of America*, vol. 101, no. 6, pp. 3722–3727, 06 1997.

[2] Verdolini, K., & Ramig, L. O. (2001). Review: Occupational risks for voice problems. Logopedics Phoniatrics Vocology, 26(1), 37–46. https://doi.org/10.1080/14015430119969

[3] “Analog and digital MEMS microphone acquisition and processing software expansion for STM32Cube” *st.com*, 2025. https://www.st.com/en/embedded-software/x-cube-memsmic1.html#overview (accessed Feb. 14, 2025).

[4] Hillenbrand, James & Cleveland, Ronald & Erickson, Robert. (1994). Acoustic Correlates of Breathy Vocal Quality. Journal of Speech, Language, and Hearing Research. 37. 769-778. 10.1044/jshr.3704.769.

[5] Association for Computing Machinery, “ACM code of ethics and professional conduct,” *Association for Computing Machinery*, 2018. https://www.acm.org/code-of-ethics

[6] IEEE, “IEEE Code of Ethics,” *ieee.org*, Jun. 2020. https://www.ieee.org/about/corporate/governance/p7-8.html

[7] Assad, J. P., Magalhães, M. de C., Santos, J. N., & Gama (2017). Vocal dose: An Integrative Literature review. Speech, Language, Hearing Sciences, and Education Journal. https://www.scielo.br/j/rcefac/a/sMbmbbjRfkWVJqzs4Xzmhgj/?lang=en&format=pdf

Appendix A: Requirement and Verification Tables

Power Subsystem

Requirements	Verification
1. The power subsystem shall be able to supply power for at least 8 hours of continuous usage.	• To verify the requirement, the device shall be running with BLE enabled for eight hours starting with full battery. • At the end of the eight hours, functionality will be verified on the receiver and battery voltage will be read.
2. The power subsystem shall be able to supply power for battery voltages from 2.0 V to 3.5V	• To verify this requirement, the battery shall be replaced with a variable voltage power supply. Functionality of the dosimeter will be tested with supply voltages of 2.0 V to 3.5 V in increments of 0.5 V.

Sensing Subsystem

Requirements	Verification
1. The sensing subsystem should pick up the user's voice.	• To verify requirement one, our device shall be worn by a user. The user will speak at a moderate volume while the dosimeter is recording.
2. The sensing subsystem should isolate external noise such that features are still preserved	• To verify this requirement, this device should be worn in a quiet room with an external speaker and measurement microphone. • Both the dosimeter and the microphone will be turned on to record. The user will speak for ten seconds while white noise is playing in the background.

	• The user will stop talking and let the dosimeter record the background noise for another ten seconds • The power of the speech vs white noise will be compared between the measurement microphone and voice dosimeter.

Bluetooth Subsystem

Requirements	Verification
1. The Bluetooth subsystem should be able to transmit real-time accelerometer data with less than 1 second of latency.	• Our device shall be worn by a user in a quiet room. A camera will be set up that can view the computer/phone screen of the receiver. The user will record a video of the screen while making a vocalization. Afterwards, the video will be viewed in a video editing software to measure latency.
2. The Bluetooth subsystem should be able to transmit at a range of at least 5 meters.	• During regular use, we will place the device that we want to connect to through Bluetooth at varying distances and record if we are still able to receive the accelerometer data. If we are able to consistently receive data for the devices being separated by at least 5 meters, then the verification is a success.

Memory Subsystem

Requirements	Verification
1. The device shall be able to write sensor data to the flash memory in real-time.	• The user shall create a timer on the MCU and measure the speed of writing 16 kilobytes of data to flash. This should take less than 1 second

	as 16 kilobytes corresponds to 1 second of sensor data.
2. The device shall be able to read sensor data from the flash memory in real-time.	• The user shall create a timer on the MCU and measure the speed of writing 16 kilobytes of data to flash. This should take less than 1 second as 16 kilobytes corresponds to 1 second of sensor data.

Processing Accuracy

Requirements	**Verification**
1. The sensing subsystem should be accurate enough to determine the sound pressure level within 2dB, fundamental frequency within 5Hz, and cepstral peak prominence within 2dB.	• To verify requirement one, our device shall be worn by a user in a quiet room. A measurement microphone will be placed one meter from the user's voice. The user will read a passage. • The three voice features will be compared from our voice dosimeter to the features using the measurement microphone and the same algorithms.

Mechanical Design

Requirements	**Verification**
1. The silicone enclosure should be comfortable.	• We will test comfort by having a sample population try on the device and state their comfort level.
2. The TPU enclosure should not impede the data transfer process.	• We will use the app to ensure that data is being transmitted by the accelerometer with the TPU enclosure.

3. The tape should hold for at least eight hours.	• We will make sure that the tape is appropriately attached and do actual trial runs to make sure that the tape stays on for at least eight hours.

Appendix B: Materials List

Description	Manufacturer	Part Number	Quantity per PCB	Cost per unit
IC REG LIN 1.8V 300MA TSOT-23-5	Texas Instruments	TPS72218DBVR	1	$1.41
IC FLASH 4GBIT SPI 83MHZ 8UPDFN	Micron	MT29F4G01ABBFDWB-IT:F	1	$3.56
DC DC CONVERTER	Texas Instruments	TPS631000DRLR	1	$1.42
AH168M	Taiyo Yuden	RF ANT BLUETOOTH/WLAN/ZIGBEE SMD	1	$4.42
CONN HEADER SMD 14POS 1.27MM	Samtec	FTSH-107-01-F-DV-K	1	$3.39
RF FILTER LOW PASS 2.45GHZ 0603	TDK	DLF162500LT-5028A1	1	$0.48
IC RF TXRX+MCU 802.15.4 48UFQFPN	STMicroelectronics	STM32WB55CGUx	1	$6.08
BATTERY RETAINER COIN 10MM SMD	MPD	BHX1-1025-SM	1	$0.87
DIGITAL VOICE VIBRATION SENSOR	Syntiant	V2S200D	1	3.93
CAP SMD 0603 47UF	Murata	GRM188R60J476ME01D	1	$0.36
CAP SMD 0603 22UF	Murata	GRM188R60J226MEA0J	1	$0.16
CAP SMD 0603 4.7UF	Murata	GRM188R61E475KE11D	4	$0.19
CAP SMD 0603 2.2UF	Murata	GRM188R61E225KA12D	2	$0.11
CAP SMD 0603 1UF	Murata	GRM188R61A105KA61D	2	$0.11
CAP SMD 0603 100NF	Murata	GCJ188R71E104KA12D	6	$0.13
CAP SMD 0603 10NF	Murata	GCM188R72A103KA37D	1	$0.11
CAP SMD 0603 100PF	Murata	GRM1885C1H101JA01D	1	$0.08
CAP SMD 0603 10PF	Murata	GCM1885C2A100JA16D	2	$0.10
CAP SMD 0603 0.8PF	Murata	GRM1885C1HR80BA01D	1	$0.13
CAP SMD 0603 0.3PF	Murata	GQM1875C2ER30BB12D	1	$0.42
RES SMD 0603 100K OHM	Panasonic	ERJ-3EKF1003V	3	$0.10
RES SMD 0603 33K OHM	Panasonic	ERA-3AEB333V	1	$0.10
RES SMD 0603 12K OHM	Panasonic	ERA-3AEB123V	1	$0.10
RES SMD 0603 100 OHM	Panasonic	ERJ-3EKF1000V	1	$0.10
IND SMD 0805 10UH	TDK	KLZ2012MHR100HTD25	1	$0.20
IND SMD 0603 1UH	Murata	DFE18SAN1R0ME0	1	$0.33
IND SMD 0603 10NH	Murata	LQG18HH10NJ00D	1	$0.18
IND SMD 0603 2.7NH	Murata	LQG18HH10NJ00D	1	$0.10
CRYSTAL SMD 32.0MHZ	NDK	NX3225SA-32.000MHZ-STD-CSR-1	1	$0.69

CRYSTAL SMD 32.768KHZ	NDK	NX2012SA-32.768K-STD-MUB-1	1	$0.82

Table 1: Bill of Materials for a single PCB

Description	**Cost**
Passive Components	$52.30
Connectors	$20.34
Integrated Circuits	$90.50
Development Boards	$86.57
Enclosure	$10.00

Table 2: Development Material Costs

Appendix C: Schematic and PCB

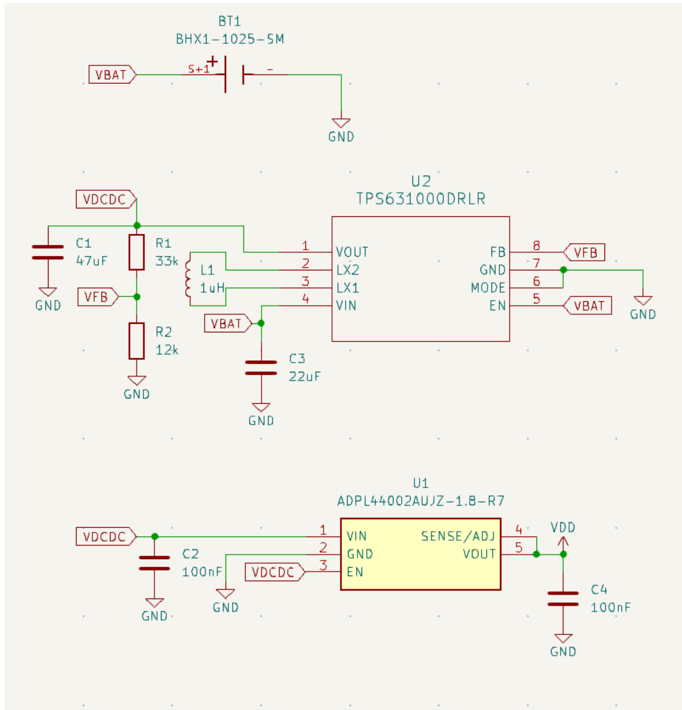

Figure 8: Power Subsystem Schematic

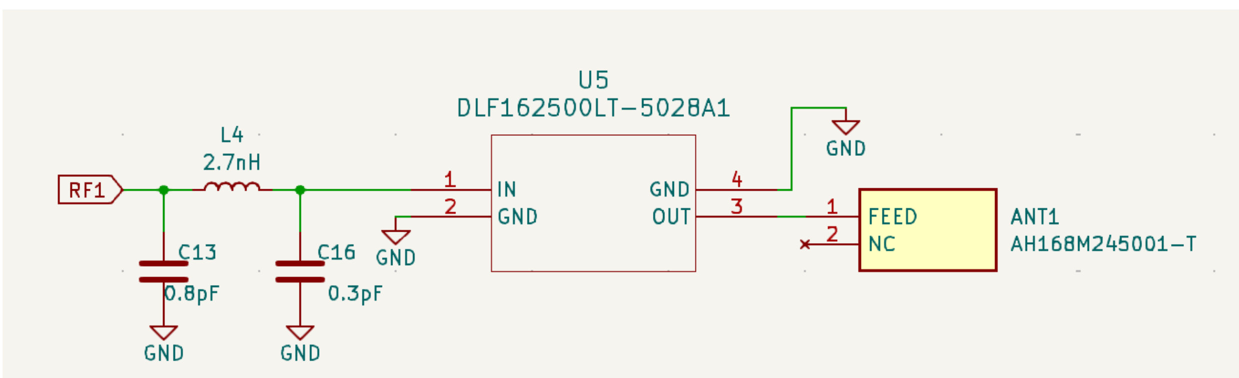

Figure 9: RF Front-end Schematic

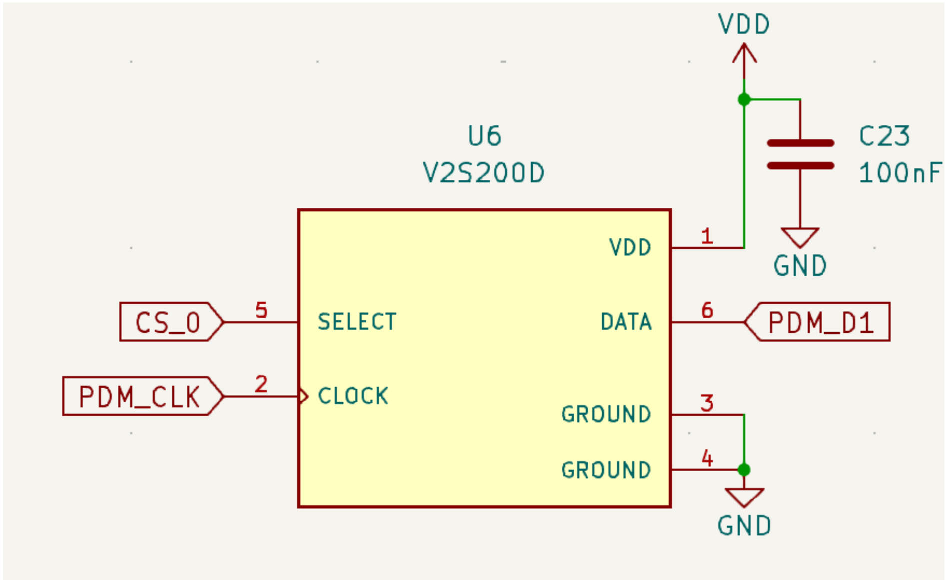

Figure 10: Vibration Sensor Schematic

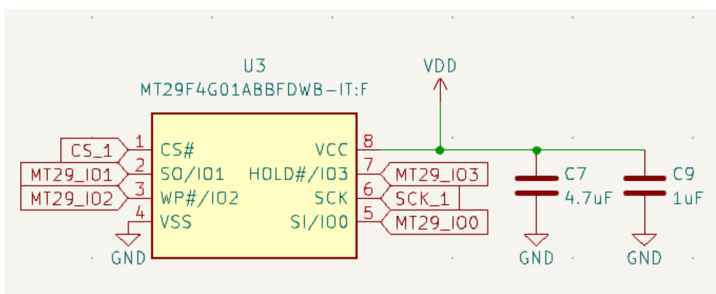

Figure 11: Memory Subsystem Schematic

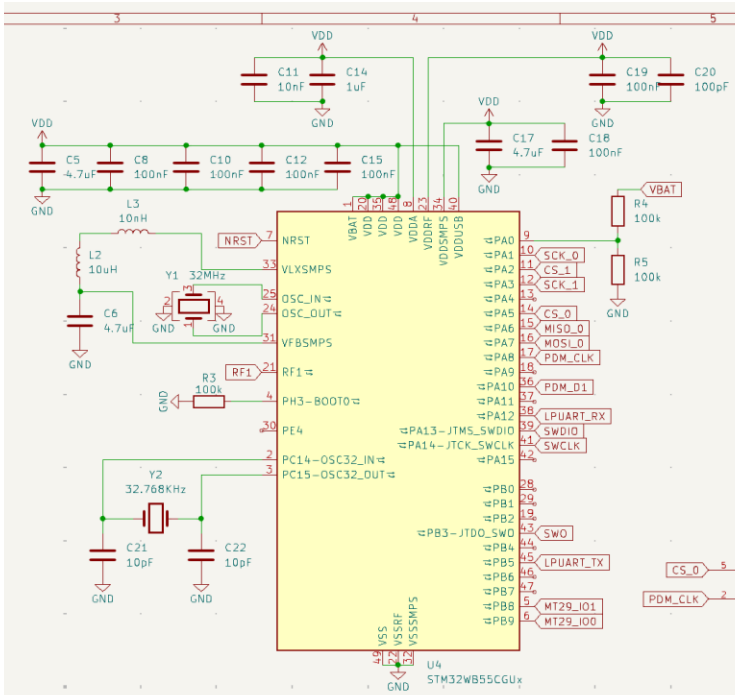

Figure 12: MCU Schematic

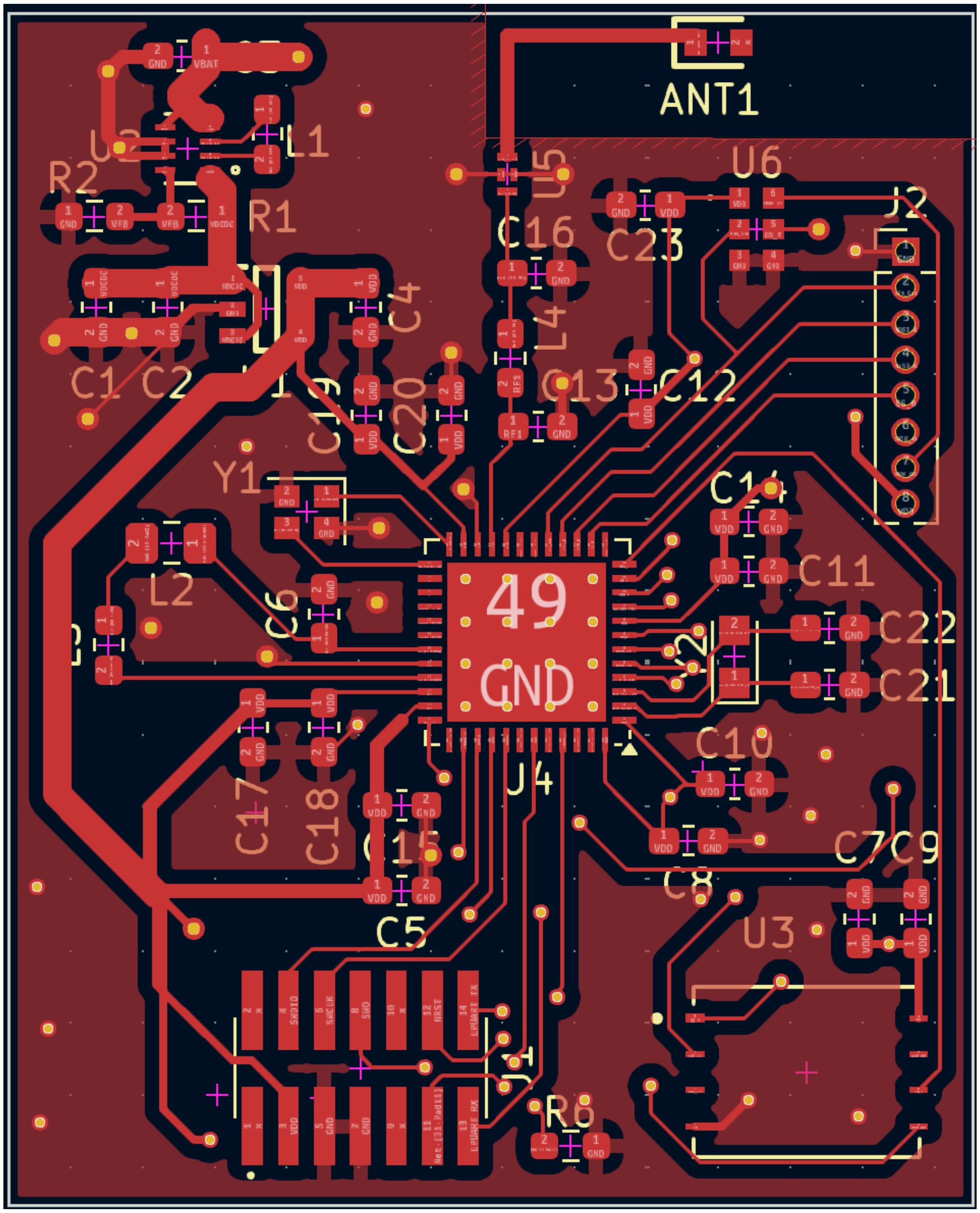

Figure 13: PCB Front

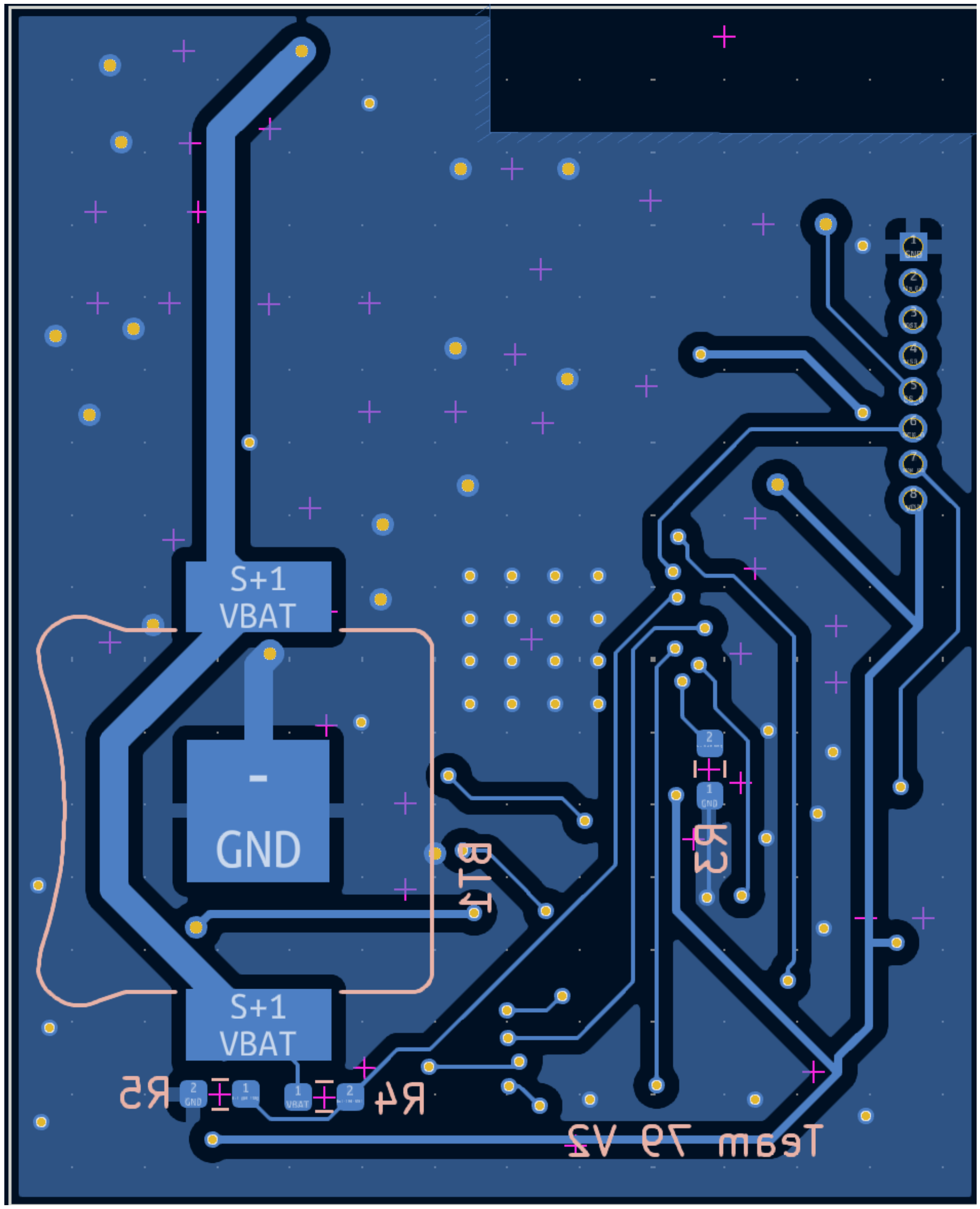

Figure 14: PCB Back

Appendix D: Software Repository

Our firmware and website code are available here:
https://github.com/rizk2/ECE445Code_VDOS