Application of the Fourier Transform to Explain Noise Suppression in Relation to the Rubik's Cube

Anderson Gao

West Point Grey Academy, Vancouver, BC, Canada

andersongaoht@gmail.com

Abstract. Digital Noise Suppression (DNS) is a mechanism that is implemented within many video conferencing apps. It makes use of the Discrete Fourier Transform (DFT), which essentially "cuts out" high and low frequencies in each sound input to achieve a comfortable quality. This study attempted to explore the DNS of Zoom, a popular video conferencing software, after experiencing an unexpected noise suppression event concerning the Rubik's Cube, which prompted an investigation to identify a threshold at which the frequencies are cut out. Using a consumer grade setup and MATLAB, a computer programing language, DFTs were performed on Rubik's Cube sounds and compared with the noise suppressed signals in order to find the interval at which DNS occurs at. This research paper was written in affiliation with the Pioneer Academics program in the summer of 2022.

1 Introduction

1.1 The Broad Scope

As the world continues to adapt to the "online era" of schooling, corporate meetings, and group gatherings, many video conferencing apps have risen to facilitate these needs. Popular interfaces such as Zoom, Microsoft Teams, and Discord have all seen a dramatic increase in users over the past two years due to the COVID-19 pandemic. However, this novel surge in video conferencing users has resulted in countless embarrassing incidents due to open microphones, poor choice of environment, or ambient noise. To combat these all too familiar problems, many software engineers have integrated noise suppression capabilities into their programs – these should not be confused with Active Noise-Cancelation, however, as both capabilities tackle different problems.

Active Noise-Cancelation (ANC) is a mechanism used by many headphone companies to give the effect of noise-cancelation. Contrary to popular belief, the headphones themselves are not completely blocking outside noise, rather, each earpiece plays a set of sound waves with opposite peaks and troughs, ensuring maximum destructive interference. On the other hand, Digital Noise Suppression (DNS), simply referred to as "Noise Suppression", is a mechanism that utilizes real-time information to selectively mute certain frequencies of sound.[1] Most commonly, this is used to cut out ambient noise and excessively high and low frequencies. Typically, a video conferencing software would receive a sound input and adjust the signal to accommodate one's voice, which is extremely helpful in loud environments without soft surfaces to absorb the excessive frequencies. DNS is also used in video and music production when manual editing is not convenient. However, for the purpose of this research paper, we only investigated the DNS capabilities of various video conferencing programs and their tolerance to different frequencies of sound.

1.2 Applications

DNS makes use of the Discrete Fourier Transform (DFT), wherein a complex signal, usually a sound wave, is separated into its component frequencies – a finite sum of sine waves. The program then mutes high and low-frequency sound waves before performing an Inverse Discrete Fourier Transform (inverse DFT) and playing the modified audio track to listeners. The majority of this research paper focuses on frequency analysis and the application of the DFT. As a result of an unexpected noise suppression event, the rest of this research paper explores a unique application.

Invented in 1974, the Rubik's Cube quickly reached international popularity with millions of people turning away at the cube.[2] However, many design flaws in the original Rubik's Cube rendered it useless in competition-style situations, where speed was a major deciding factor in the winning solve. To combat this, many third-party companies started to develop their own "speed cubes", which were much easier to turn and allowed room for misalignment errors – a design feature that would eventually cause the original Rubik's Cube to fall out of favor among the top speedcubers at the time.

As different speed cubes slowly developed into their own brand, each company settled on a "formula" for their own cube, including, but not limited to, type of plastic, turning mechanism, and core design. Each unique feature contributed to the cube's overall "feel" and, in turn, produced a characteristic sound. For instance, a company's cube may sound "crunchy", while its competitor's cube may seem "buttery". The frequency sound representations of these cubes vary immensely, which is what this research paper investigated.

As mentioned beforehand, this investigation was prompted by the unexpected noise suppression of a Rubik's Cube while performing an unrelated demonstration on Zoom – the cube was unintentionally muted. Amidst this frustration came the idea to write this research paper to find the interval at which noise suppression occurs for Rubik's Cubes, an aspect that is elaborated on in Section 2.1.

By experimenting with the DNS systems of various video conferencing applications, we will be able to determine an optimal cube turning volume and speed, as well as find the optimal speed cube brand to remain undetected by the DNS systems of these video conferencing applications.

2 Methodology

2.1 The Discrete Fourier Transform

As alluded to in Section 1.2, the Fourier Transform is a mathematical operation used to decompose a function into its fundamental sinusoidals. Equation 1 shows the continuous Fourier Transform of the function, $f(t)$, commonly denoted with the symbol $\mathfrak{F}$.[3]

$$\mathfrak{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{-i2\pi kt} dt \tag{1}$$

However, in order for the above Fourier Transform to be applied, the function must be continuous due to the properties of integrals – a restriction that severely diminishes its applicability in the real world. As mentioned in Section 1.2, the Discrete Fourier Transform (DFT) serves as a mediator between the perfect world of mathematics and the imperfect nature of reality by allowing discrete inputs of *N* data points with *k* as the current frequency.

$$X_k = \sum_{n=0}^{N-1} X_n \cdot e^{-i2\pi k\frac{n}{N}} \tag{2}$$

Equation 2 can also be separated into real and complex parts using the modified Euler's Identity:

$$e^{-ix} = \cos x - i \sin x \tag{3}$$

This is used to derive an alternate expression for the DFT.[4]

$$X_k = \sum_{n=0}^{N-1} X_n \cdot [\cos\left(2\pi k \frac{n}{N}\right) - i \sin\left(2\pi k \frac{n}{N}\right)] \tag{4}$$

Using this method in conjunction with MATLAB – an analytical programming language for complex calculations – we can use the `wav_fft` function to apply a transform onto the raw signal from the Rubik's Cubes. It is also important to acknowledge that the `wav_fft` function is considered a Fast Fourier Transform (FFT) and not the traditional DFT, however, this paper utilizes both terms interchangeably as the FFT is merely a method of calculating the DFT. After obtaining the FFT, one can look for distinguished and large spikes in the graph to find the main frequencies.[5]

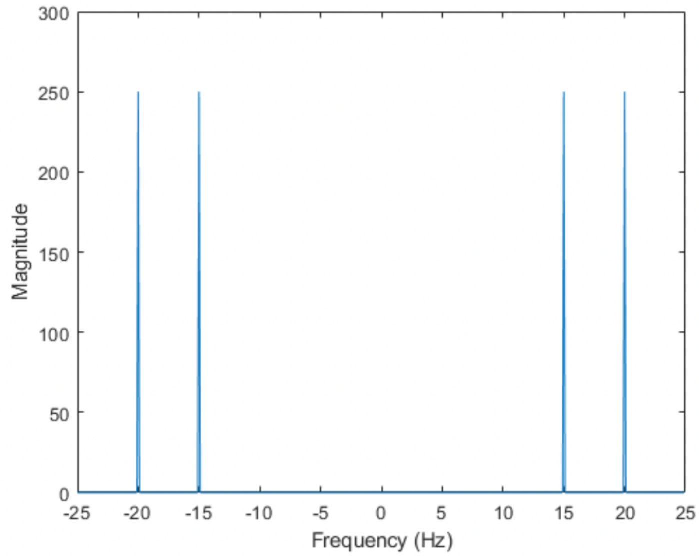

Fig. 1: Example of a clean Fourier Transform

Figure 1 is a good example of an informative Fourier Transform. The spikes in magnitude indicate that the original signal contained components of 15 Hz and 20 Hz – the negative frequencies should be disregarded as they are not applicable in the real world and are a result of the complex number in Eq. 1. Using the same logic, applying a FFT to the raw sound signal of the Rubik's Cubes can give us valuable insight of each cube's component frequencies. This then allows analysis to find a frequency "threshold" at which the noise suppression activates and "cuts out" certain frequencies.

2.2 Data Collection

Three different cubes were selected for this experiment: the GAN 356 Air M, QiYi Valk M, and YuXin Little Magic M. Each cube was selected for its unique "texture" and subsequent availability. All cubes used in this experiment would be categorized into the standard 3x3 Rubik's Cube.

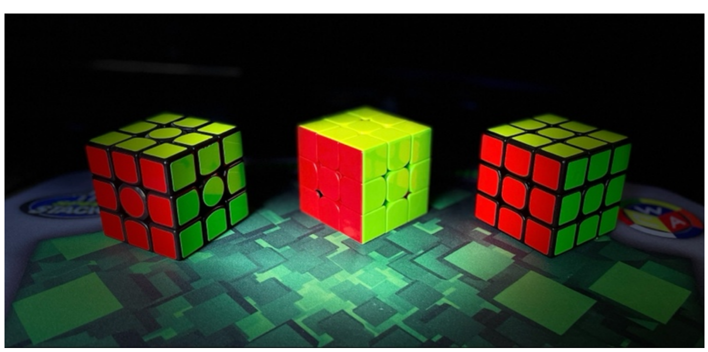

Fig. 2: The three cubes from left to right: GAN 356 Air M, QiYi Valk M, YuXin Little Magic M.

As alluded to in Section 1.2, the turning sounds created by these cubes were unexpectedly suppressed through Zoom – a completely unintended side effect from noise suppression technology. Thus, a question was proposed: Does the noise suppression have a specific threshold of frequencies with which it cuts out?

To answer this question, all three cubes were recorded with a Blue Yeti Microphone into a software called Audacity. This allowed the audio files to be exported in the `.wav` format, which were then graphed into MATLAB via the `audioread` function. Figure 3 shows the audio graph produced by the GAN 356 Air M.

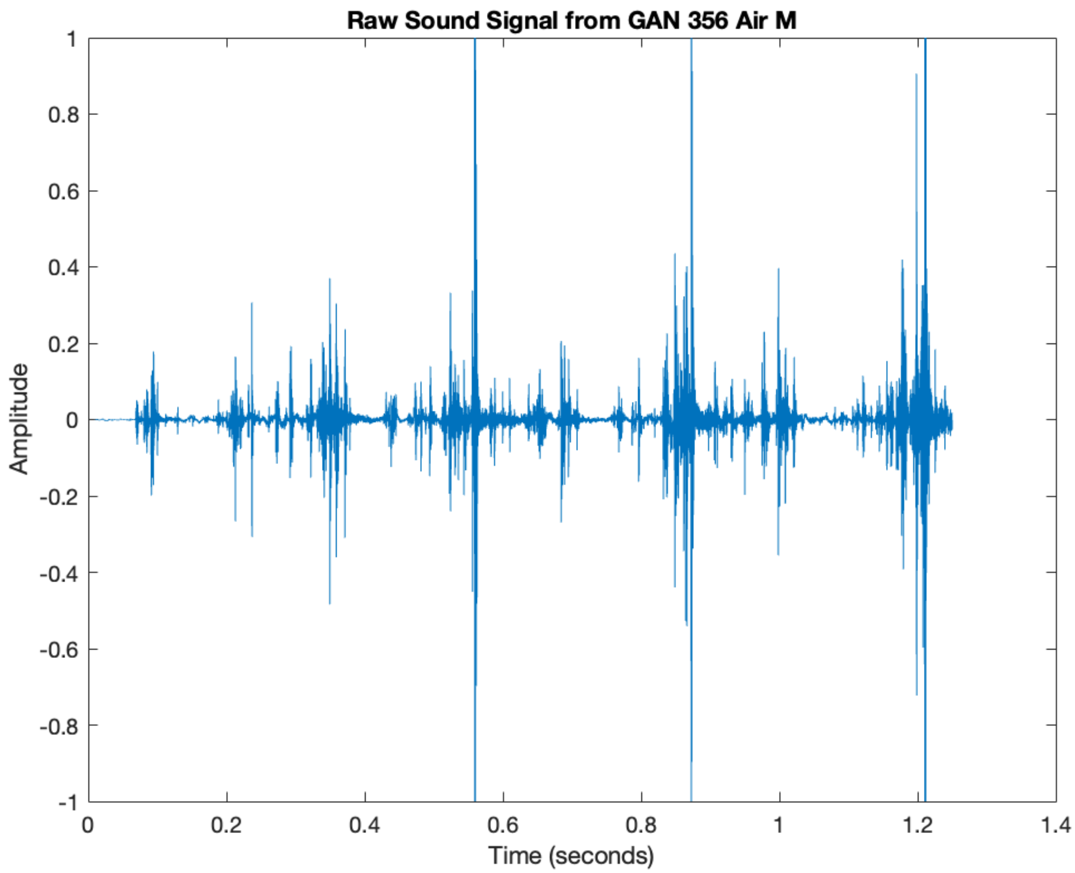

Fig. 3: A raw sound graph produced by MATLAB of the GAN 356 Air M.

Graphs of the QiYi Valk M and YuXin Little Magic M are available in Section 3.1.

Following the raw recordings, the same cubes were then recorded into Zoom to see the changes that would occur. For this experiment to run smoothly, two computer setups were used: a 2020 MacBook Pro 13-inch (factory microphone settings) was set up in one room, while a desktop computer was running in another room. Both computers were in the same Zoom call to allow audio exchange and each Rubik's Cube sound was recorded into the MacBook Pro.

These sounds would then be played in the second room through the built-in speakers of the LG 34UC98 monitor into a Blue Yeti Microphone and inputted into Audacity. This was again converted into a `.wav` sound file and read into MATLAB. Figure 4 only shows the graph for the GAN 356 Air M as an example, and the Zoom outputs of the other two cubes can be found in Section 3.1.

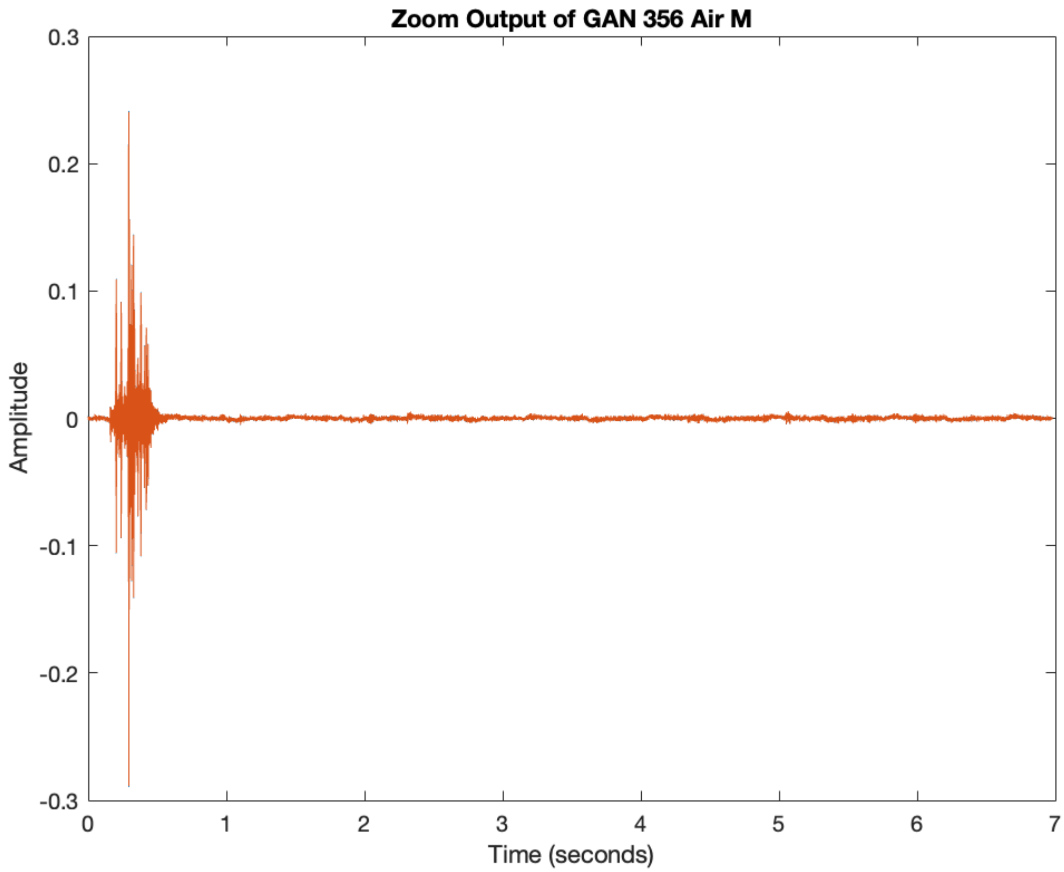

Fig. 4: The signal produced by the GAN 356 Air M after noise suppression by Zoom.

As evident from the graph, the sound is "cut off" after the first ~0.5s, which suggests evidence of noise suppression in use. This characteristic flatline seems to be a common occurrence with repetitive ambient noises – a useful point of consideration. Section 3.2 puts forth a hypothesis to explain this phenomenon, however, further analysis is required.

3 Results

3.1 QiYi Valk M and YuXin Little Magic M

As mentioned in Section 2.2, the raw sound signals produced by the QiYi Valk M and YuXin Little Magic M are as follows:

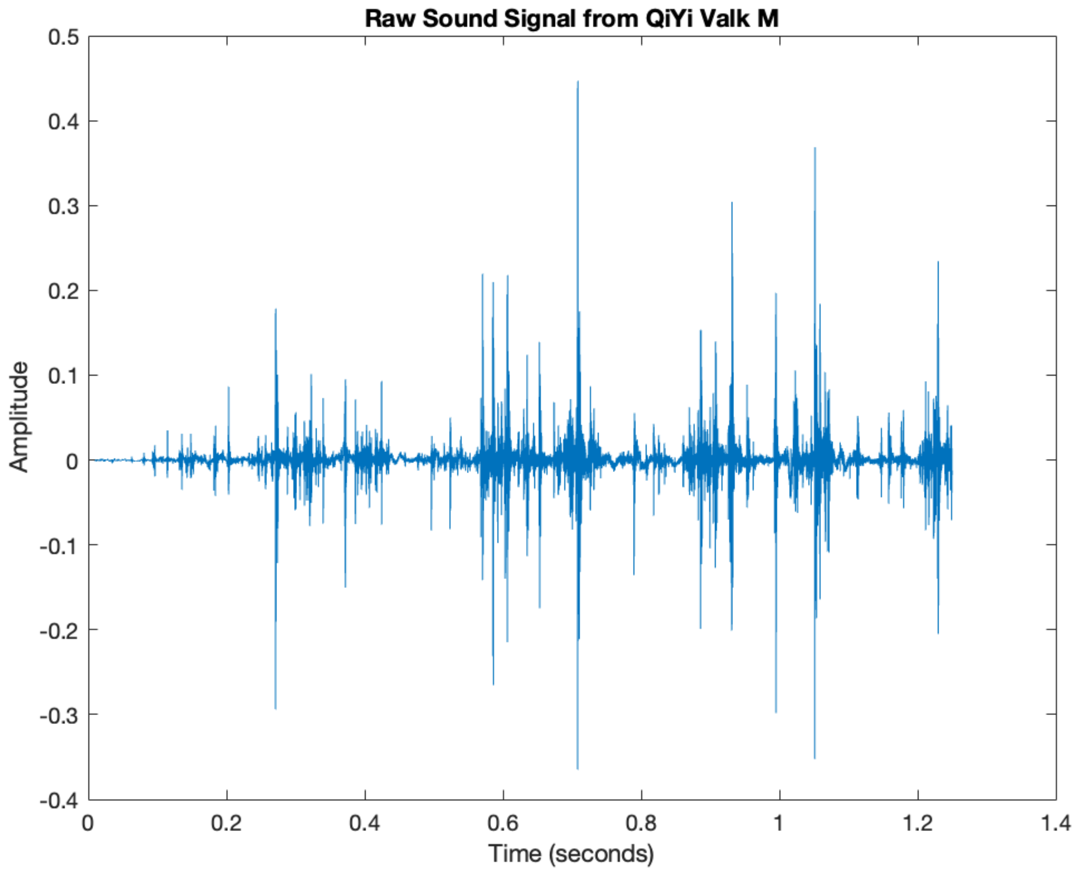

Fig. 5: A raw sound graph produced by MATLAB of the QiYi Valk M.

Comparing this with Fig. 2, it is evident that the QiYi Valk M has a much less "aggressive" tone compared to the GAN 356 Air M due to the weaker peaks in the graph. This, however, did not seem to cause much of a change when put through the Zoom DNS.

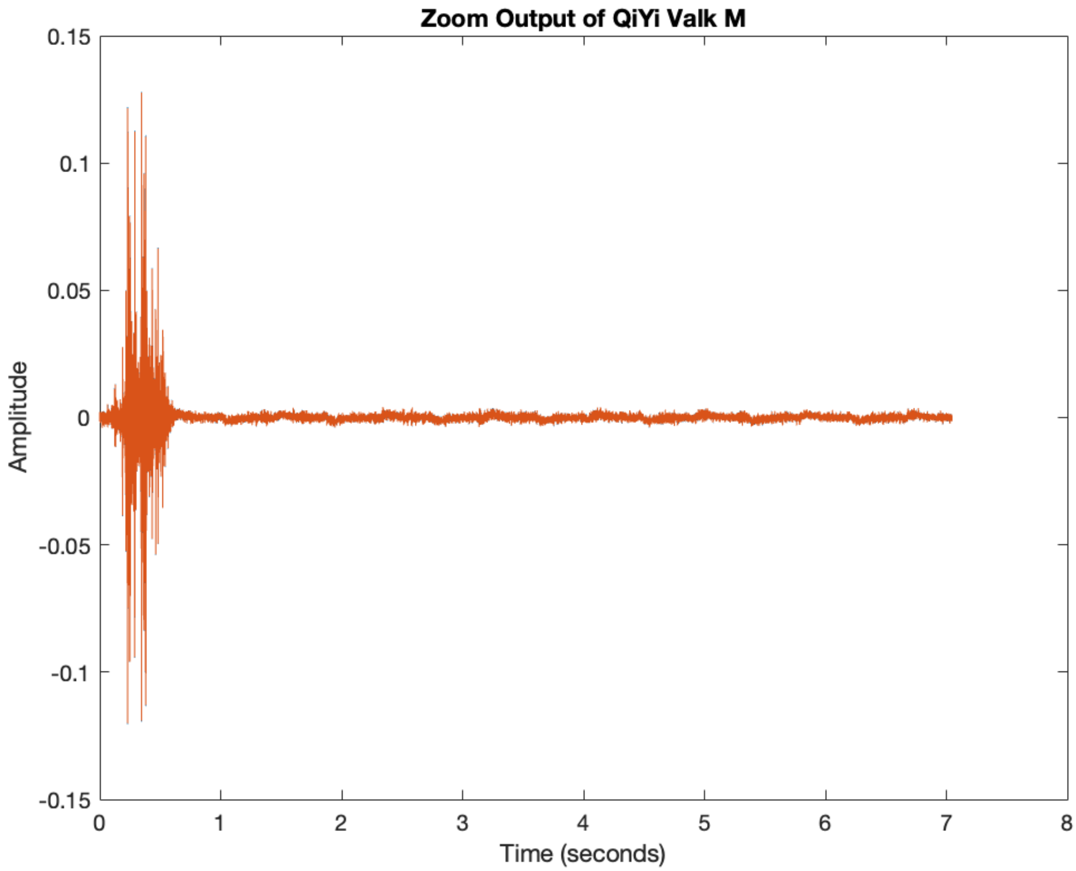

Fig. 6: The signal produced by the QiYi Valk M after noise suppression by Zoom.

It is evident from the graph that despite a much more "mellow" raw signal, the Zoom DNS continues to cut out the sounds of the QiYi Valk M after ~0.7s. However, one slight difference is the length that the signal prevails; the GAN 356 Air M flatlines slightly earlier than the QiYi Valk M, which could be due to the strong peaks that are not prevalent in the QiYi Valk M – a point of consideration. Additionally, the YuXin Little Magic M also yielded some interesting results.

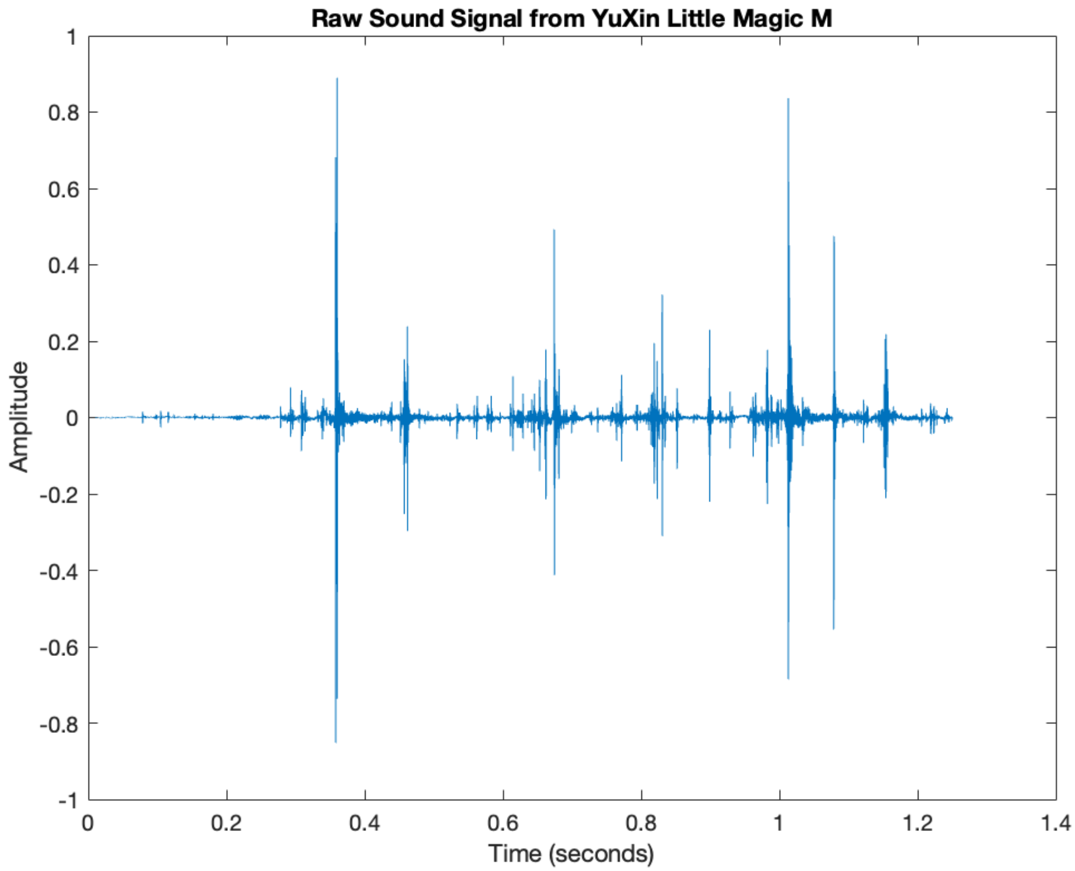

Fig. 7: A raw sound graph produced by MATLAB of the YuXin Little Magic M.

Comparing the YuXin Little Magic M to the likes of the GAN 356 Air M and QiYi Valk M, we see a much more mild characteristic with thin and relatively low peaks. We also see a strange new phenomenon when looking at the signal received when passed through the Zoom DNS.

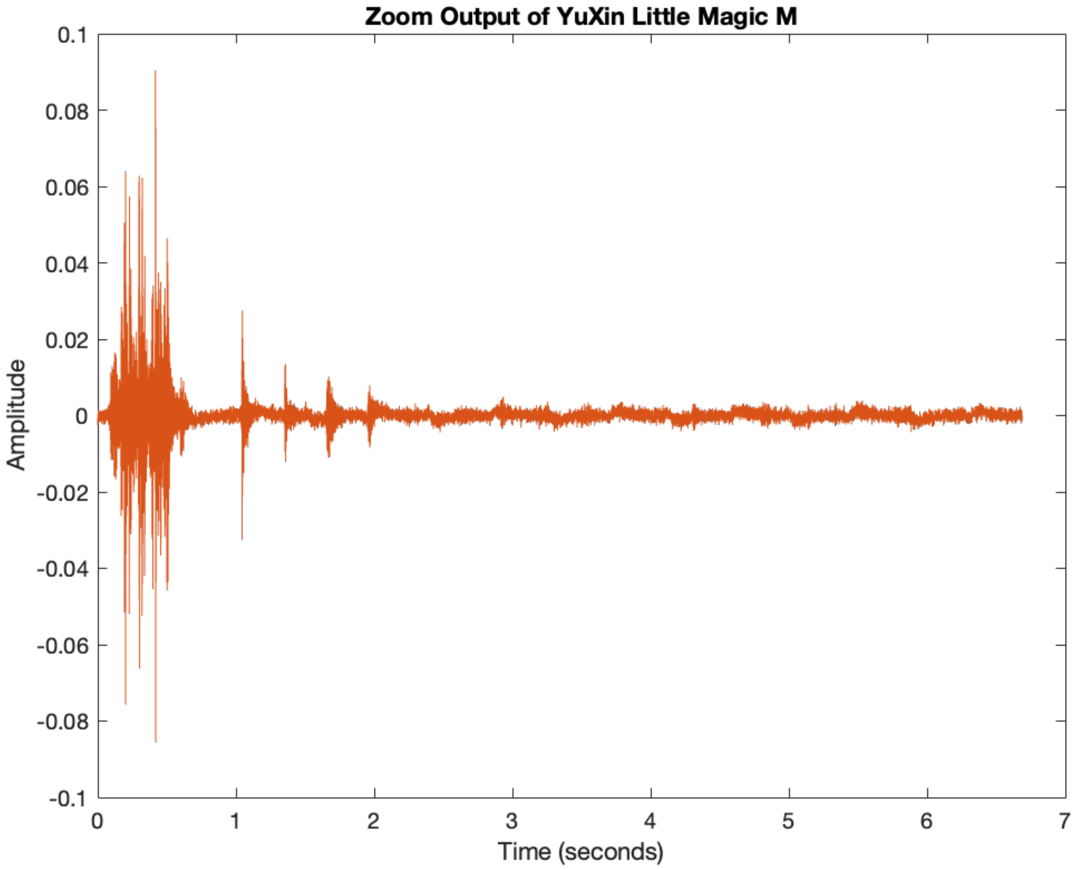

Fig. 8: The signal produced by the YuXin Little Magic M after noise suppression by Zoom.

An interesting anomaly that is only present in the sound signal from the YuXin Little Magic M is the presence of four distinct "spikes" that gradually decay over a constant time interval into the regular flatline. This is certainly a phenomenon to discuss further and provides insight into the "muting" threshold of the Zoom DNS. Section 3.3 poses a hypothesis for these strange spikes.

The amplitude of each spike is also smaller than the GAN 356 Air M and QiYi Valk M, but this is expected as the YuXin Little Magic M did have a qualitatively quieter sound when being recorded.

3.2 Analyzing the FFT

To reach a meaningful explanation for the frequency threshold of the Zoom DNS, one can analyze the FFT of the raw sound signals from the Rubik's Cubes to find the major frequencies in a given sound.

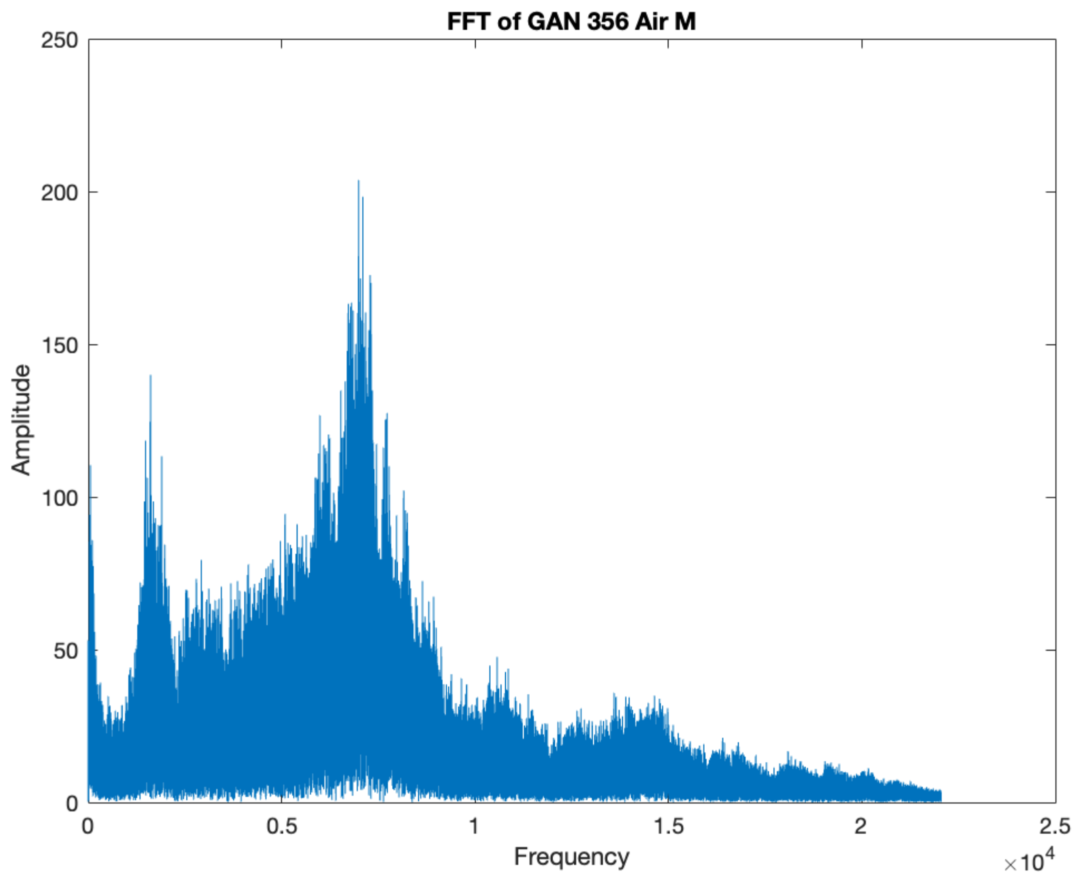

Fig. 9: The Fast Fourier Transform produced by MATLAB of the GAN 356 Air M.

With all non-musical and real world applications, this Fourier Transform is certainly not as clean as the one in Fig. 1 due to the complex nature of the sound of the Rubik's Cube. Nonetheless, this FFT still gives us some insight into the component frequencies concerning the GAN 356 Air M; despite the inevitable noise, there is still a distinct spike in amplitude at ~6,700 Hz, meaning a major component frequency was ~6,700 Hz.

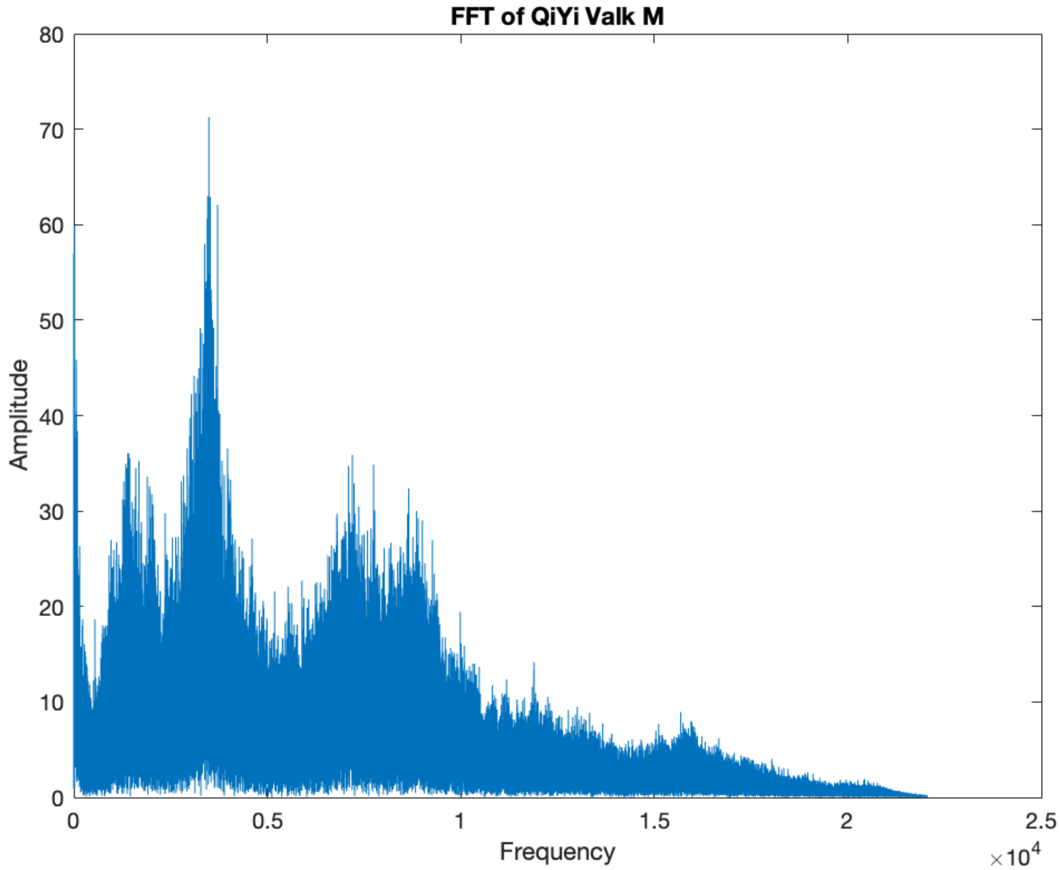

Fig. 10: The Fast Fourier Transform produced by MATLAB of the QiYi Valk M.

Looking at the FFT of the QiYi Valk M, there is a clear spike at the frequency of ~4,000 Hz, indicating that a major component frequency is around 4,000 Hz. With this information, it is evident that the Zoom DNS mutes frequencies between ~4,000 Hz and ~6,700 Hz.

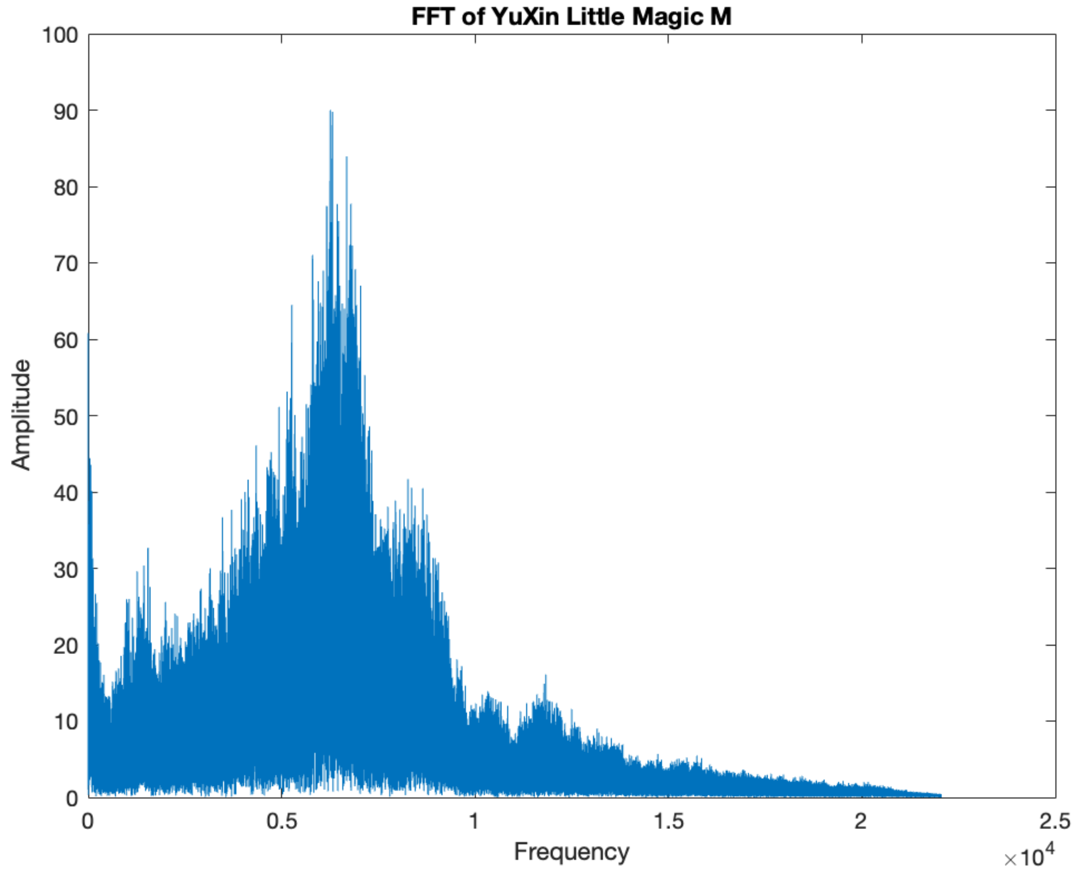

Fig. 11: The Fast Fourier Transform produced by MATLAB of the YuXin Little Magic M.

Lastly, the FFT of the YuXin Little Magic M shows a distince spike in amplitude at ~6,000 Hz. This is an interesting number as the data from the GAN 356 Air M and QiYi Valk M show that the Zoom DNS would completely suppress frequencies between ~4,000 Hz and ~6,700 Hz, yet according to Fig. 8, this is not the case as the YuXin Little Magic M contains the four distinct spikes before flatlining.

3.3 An Explanation for the Incomplete DNS on the YuXin Little Magic M

The only possible explanation to describe the four decaying spikes in Fig. 8 would be that between 4,000 Hz and 6,700 Hz, there is an interval where the DNS system is unsure of what to do; initially, the DNS detects that the signal is within the "muting" threshold, thus the frequency is suppressed. However, the system realizes that this frequency may be classified into another category of important sounds, and in turn, plays it through Zoom. This may go on for a while, until the signal ceases.

However, given the elaborate explanations, there is also a high possibility that the four spikes in amplitude were just a product of hardware malfunction – a theme that is discussed in Section 4.1.

4 Conclusion

4.1 Possible Sources of Error

As alluded to in Section 3.3, many confounding factors could have contributed to the quality of the data collected. For instance, this experiment was conducted with consumer grade electronics and quite frankly, a rudimentary setup. The built-in microphone on the MacBook Pro 13-inch may have unintentionally contributed to the noise suppression, causing an inaccurate Zoom graph. It may have also malfunctioned during the recording session, resulting in the unusual spikes seen in Fig. 8. The microphone quality may have also confounded the readings, but this is merely a hardware limitation.

Likewise, the other end of the recording setup may have encountered similar problems as the microphone. For example, the built-in monitor speakers that the Zoom audio was being play from could have also confounded the data as the LG 34UC98 speakers are not the best instruments for producing true sound. This can also be extended to the Blue Yeti Microphone, which could also contribute to the confounding of sound signals due to hardware limitations. All in all, there are many improvements that could be made to this experiment design in order to obtain more reliable and valuable data.

4.2 Areas for Future Research

This research study was conducted under many time constraints and technological restrictions. If given more time, this study would have conducted more trials to rule out the possibility of a hardware malfunction – a point of advice for future research papers. In addition, future improvements could involve testing different turning "styles" as this study only experimented with a systematic repetition of turns, which is far from the convention of a normal solve. As a result of this, we are unable to conclude whether the noise suppression is truly cutting out frequencies because it is outside of the "safe interval", or if it is just a function of the repetitive nature.

In conclusion, the Zoom DNS is much more complex than initially thought of, and hopefully future research can fully dissect and analyze the inner workings of the system to a greater depth.

Acknowledgements

I would like to thank Professor Arthur Western for his mentorship throughout the research process. This research paper was written in affiliation with Pioneer Academics.

References

1. J. Kruckow, Sensear (2017).
2. H. Reese, Smithsonian.com (2020).
3. J.F. James, A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering (Cambridge University Press, Cambridge, 2015).
4. S. Smith, Digital Signal Processing: A Practical Guide for Engineers and Scientists (Newnes, Amsterdam, 2003).
5. T. Yoshizawa, S. Hirobayashi, and T. Misawa, SpringerOpen (2011).