NONLINEAR MIXING SYSTEM

Hairul Hafizi bin Hasnan

Doctor of Philosophy

University of York

Music

September 2020

Abstract

This thesis presents a novel massively mutual compression system for use as a nonlinear mixing tool. In this interdisciplinary work, the system is explained and explored from the perspectives of both audio engineering and music production aesthetics. First, the traditional method for presenting the time-frequency energy distribution of musical audio signals, the short-time Fourier transform, is assessed. Following that, an alternative, the constant-Q transform, is suggested as a better approach to illustrating variations in time-frequency content of audio signals.

By exploring the architecture of digital compressors, a method to create a bilateral cross-adaptive architecture is proposed and created in the Max/MSP visual programming language. This platform is chosen because of its routing flexibility and performance. In this thesis, it was found that commercially available digital audio workstation applications are not able to support the feedback cross-modulation which is crucial for this architecture. The behaviour of the system, in particular a 'double knee' curve, is characterised, verified, and explained by a series of tests.

Then, two original audio examples are presented which are mixed in different ways to highlight how the system developed in this research can be utilised. These examples highlight some unique properties of this system. The first is the lack of one perpetually dominant element in the mix. Secondly, constant changes are introduced into the tonal balance of the mix. The research also found that the 'top down' approach to mixing is the most intuitive approach to using the system.

Finally, user tests were conducted involving 15 participants. The experiment found that users will most likely use the architecture as a mixing tool to reduce masking. Furthermore, most of the participants agree that there are advantages of utilising this tool in ther production. Additionally, a majority of the respondents have discovered mixing technique that are unique to the nonlinear mixing system.

List of Contents

3 Novel Architecture and Analysis

List of Tables

List of Figures

List of Accompanying Materials

Audio Files

AudioFile1.wav	Stereo sound file of ambient music mixed using the nonlinear system. The parameters were set to extreme values.
AudioFile2.wav	Stereo sound file of ambient music mixed using a linear system.
AudioFile3.wav	Stereo sound file of ambient music mixed using a sidechain system.
AudioFile4.wav	Bass drum mixed using the nonlinear mixing system.
AudioFile5.wav	Gong mixed using the nonlinear mixing system.
AudioFile6.wav	Stereo sound file of ambient music mixed using the nonlinear system. The parameters were set to produce subtle compression.
AudioFile7.wav	Stereo sound file of experimental/hip hop music mixed using the nonlinear system.
AudioFile8.wav	Stereo sound file of experimental/hip hop music mixed using a linear system.

AudioFile9.wav	Stereo sound file of experimental/hip hop music mixed using a sidechain system.
AudioFile10.wav	Stereo sound file of experimental/hip hop music mixed using the multi-band compression system (same-band).
AudioFile11.wav	Stereo sound file of experimental/hip hop music mixed using the multi-band compression system (cross-band).

Max Patch

To execute these patches, users must install Max 8 software to their computer. It can be downloaded from the developer's website.*

NonlinearSystem.maxpat

Max patch for the eight-track nonlinear mixing system.

MultibandSystem.maxpat

Max patch for the same-band multi-band nonlinear mixing system.

CrossMultibandSystem.maxpat

Max patch for the cross-band multiband nonlinear mixing system.

* https://cycling74.com/downloads

Declaration

I hereby declare that this thesis is entirely my own work and all contributions from outside sources have been explicitly stated and referenced.

Although not forming the main body of work presented here, reference is made in the text to previously presented research carried out during the preliminary stages of work for this thesis. These publications are listed as follows:

1. Hairul H. B. Hasnan and Jeremy J. Wells, "Investigating the Behaviour of a Recursive Mutual Compression System in a Two-Track Environment," *Proceedings of the Audio Engineering Society 146th International Convention* (Dublin, 2019).
2. Hairul H. B. Hasnan, "Nonlinear Mixing," *Art of Record Production Conference 2019,* (Berklee College of Music, 2019).

1 Introduction

In music, control of dynamics is one of the key elements of expression. Dynamics generally refers to the variation of loudness in music as it is performed, yet it also effects timbre. Traditionally, composers use dynamic markings to indicate whether a section in a piece of music is played loudly or quietly. They can be used for an entire section of an orchestra, a phrase in a composition, or the evolution of an individual note. Thus, the volume at which music is performed and heard depends on decisions made by the composer, the performer's interpretation of a composition and the acoustics of a venue. However, now that music can be recorded, stored and played back in other media, the volume level is no longer confined to the aforementioned conditions. It can be changed either during recording or in post-production using a myriad of audio processors available in the studio, as well as during playback by the end user to match audio level (be it acoustic, electric or digital) to the environment or the transmission channel.

In the early years of music production, before multi-track recording became common, dynamics were manipulated manually. For example, during a recording session the engineer or producer may have asked the musicians to play louder or more quietly. Another way of altering the level of a signal was by adjusting the distance of the microphone to the source of audio. This evolved when multi-track recording became possible. In the mixing stage, faders may have been used by engineers to increase or decrease the level of recorded material. This method provided possibilities for variations of dynamics in a mix, whereby a signal that was typically played loudly during recording, for example an electric guitar with the distortion pedal activated, could be reduced in level to alleviate masking and therefore provide space for the vocals. However, as time has progressed, the process of level control has been increasingly automated or assisted by machines using audio processors that specifically address the dynamic range of a signal. These are known as compressors.

The compressor is a commonly used audio processor and has been the subject of discussion and development for a number of years. Areas of investigation around the compressor range from its architecture (e.g., analogue or digital, feedforward or feedback) to the study of its use, or an understanding of how it affects audio signals. Recently, there has been controversy around whether or not there has been excessive use of compression to make one song louder than another. This work, however, is not concerned with that debate. Instead, it describes the conception and development of a novel system for automatic dynamic range control. This system has a variety of applications from simplifying and automating the mixing process to creating unique interactions between individual mix elements that could be used for creative purposes.

Compressors, as suggested by their name, are used to reduce the dynamic range of an audio signal. However, there are a number of ways that compressors can be used to reintroduce amplitude variations in a recorded track. One approach is to use a slow attack time that allows the unprocessed signal to pass before it is compressed. This is typically used to enhance ‘punch’ for drums or percussion. Another technique is sidechain compression. Sidechain compression, also known as ducking, is commonly used in broadcasting. Historically, the main purpose of this technique has been to automate the process of attenuating the volume of the background music when the DJ talks.

Nowadays sidechain compression is also applied creatively in music production contexts to produce dynamic variations in a composition. A sustained pad, for example, has little dynamic range. However, by utilising sidechain compression, amplitude variations can be introduced on a pulse level. Of course, quite often the overall dynamic range of the mix does not change dramatically when sidechaining is implemented.

To understand this, consider a mix that consists of a constant pad sidechained to a kick drum. The pad on its own will have an increased dynamic range. If the kick drum and pad are combined without using sidechain compression, then the overall audio level will go up when the kick drum is introduced. If the kick drum and sidechained pad are combined, then when the kick drum is heard the level of the pad will be reduced. Therefore, the two in combination gives a reduced dynamic range.

The system developed for this study takes sidechain compression, as described above, one step further by implementing a completely novel mutual sidechain architecture. It is created specifically for the purpose of post-production. To the best of this author's knowledge, such a system has never been implemented prior to this research for reasons that will be explained in the coming chapters. Consequently, the outcome of this research is the production of a new tool that producers can utilise to shape the timbral or spectral content of their compositions and to assist their mixing process.

1.1 Methodology

This is an interdisciplinary piece of research; it combines audio effect design with studio-based practice and an assessment of the musicology of production, specifically the meanings and the implications of the use of automatic dynamic range compression in popular music.

In recent times, the common creative implementation of the dynamic range processor in music production is the use of sidechain compression. This is done by imparting rhythmic undulations onto various chosen elements in a mix, triggered by one source. The purpose of this research is to develop a new tool that is an evolution of that technique. Therefore, a complete understanding of the artistic use of the compressor, from the inception of its creative application to its current application, is needed. This can be accomplished by studying the following aspects: what is a compressor and how can it be applied creatively?

There are three theoretical aspects of this research. The first one involved conducting a comprehensive exploration of compression including its history, architecture and creative application in music production. By understanding its architecture this research reveals that the sidechain processing algorithm can be expanded to create a novel system. The other theoretical aspect of this study is the decision-making process for choosing the analytical method to be used. It is important to select the most suitable tool to analyse and reveal the effect of interconnected compression system to audio signals. After the theoretical groundwork has been established, inferences regarding the characteristics of the system can be made to verify a successful implementation of the algorithm. Another theoretical framework is the research design for the data sampling process on the usefulness of the system developed in this research.

For the practical aspects of this study, attempts to actualise this architecture were first made by the author using available resources at the University of York. In particular, these resources included an SSL Duality console, a number of digital audio workstations and the use of a programming language to create a new mixing system. A few tests were then conducted to verify the success of the implementation of the architecture. Once the system was correctly producing the expected characteristics, the work described in the third chapter took place. In this stage, the system was utilised in an original production work to explore its potential in music production. The outcome of this mixing process is then analysed using a number of different approaches which will be discussed in the coming chapters.

1.2 Personal Motivation

As a drummer, composer (or, as colloquially known, a beat-maker), performing musician and recording engineer, I have always been interested in groove, dynamics and music production. Over the years I have worked with compression and I have gained experience in its application. By nature, drums have a broad dynamic range. I have been told to play more quietly (or louder) during recording sessions and dynamics signal processing helps with controlling the dynamics and attenuating drum bleed (amongst other uses).

Part of the inspiration for this research was from my master's dissertation where I developed a method to fully automate a multi-track dynamics processor (specifically noise gate) to remove the noise from a multi-track recorded instrument (specifically drums) in Reaper. The inspiration came during a drum track mixing session when someone asked if there is a way to automatically gate the bleed between different parts of the kit. The advent of the creative use of compression in music production, particularly in electronic dance music sparked my interest to explore how this can be taken further.

Based on the knowledge that I have gained during my tertiary education, which was in physics with electronics and music, I would consider myself knowledgeable with regard to compressors and their applications. This knowledge was further solidified during a four-year stint working in a radio station as an in-house drummer, audio engineer and programme producer. Prior to the start of the production work in the third chapter, a few predictions on the outcome of a massively mutual sidechain compression system were made by the author and the supervisor of this project. This includes the idea of amplitude death, where oscillations in coupled components will cease because of the interaction between the systems. Proving whether such behaviour can be observed in music signal processing became one of the core research themes.

It is my hope that the outcome from this research will provide a new creative tool at the producer's disposal. This may then facilitate in creating work that is unique and encourage novel approaches to composition, mixing and production.

1.3 Chapter Overview

Chapter 1 outlines the theoretical groundwork that thoroughly explores compression, from its architecture to its creative application. It draws together a combination of ideas and techniques from multiple disciplines.

First, an alternative computational method to plot the spectral content of a song is investigated. This is to establish a method for clearly depicting song dynamics and how they can be controlled by processing tools. This method is then used throughout the thesis to demonstrate actions of the different dynamic range processes that are discussed. The reason for this is that the most common method of analysing the spectral content of a signal is less adequate compared to the approach proposed in this research in ways that will be explained. Examples created using synthesisers are then analysed using both the common approach and the method that was proposed and compared to illustrate the differences. After it is determined that one approach is superior, the rest of the section analyses the spectral content of a handful of published musical works. This serves as an exercise regarding features to look for when conducting analysis towards the end of this thesis.

Once an analytical framework has been established, the second section explores the history of compressors and the architectures that it was based on. This introduces the inner workings of compressors and supports the reader's understanding of how the architecture can be expanded to create a bilateral mutual compression and, eventually, a nonlinear mixing system.[1]

Chapter 2 gives a complete overview of the proposed system and its behaviour. The first section of the chapter describes its novel compression architecture using mathematical representations and diagrams displaying the signal flow of the system. Based on this information, a few predictions are made on the behaviour of the system that are used in the subsequent sections as the criteria to verify the successful implementation of a bilateral compression architecture.

The second section of this chapter focuses on the implementation of the bilateral compression architecture described in the first section using two chosen digital audio workstations. However, the result produced by this exercise did not conform with what was anticipated, particularly with the compression curve and frequency content of the processed signals. This led to the third section of this chapter. In this section, the bilateral compression architecture is built using a visual programming language called Max 8. Through testing, it was shown that the compression curve and the frequency content of the output correlate with the predictions, proving successful implementation of the system.

The fourth and the fifth sections explore the characteristics of the system extensively by using different parameters and types of test signals. Not only did this provide compression curves for

[1] The decision to use the term nonlinear mixing system for this research is explained on page 52.

different conditions, but it also demonstrated how the attack and release time constants affect the processing. The sixth and the seventh sections expand upon the bilateral compression system by creating a nonlinear mixing system and a multi-band nonlinear bilateral compression system. The use of these architectures is investigated in the following chapter.

Chapter 3 demonstrates how the nonlinear mixing system can be applied in music production as a tool for corrective and creative purposes. This facilitated an exploration of the system's unknown potential. It is conducted by utilising the system developed in this research to mix two original works. One of the works was mixed using five different systems (linear mix, sidechain mix, nonlinear mix and multi-band compression mix, cross-band multi-band compression mix) and two settings (extreme and subtle) to facilitate aural comparisons and highlight how different systems and settings affects the output. The results were analysed using methods that are suitable for the desired type of information.

Finally, a small-scale study was conducted to understand how users would implement this tool in their mixing environment. This was done by conducting a user test and a mix of structured and semi-structured interviews. The information obtained from the interviews was then analysed.

The final section of this thesis provides a recapitulation and a summary of the previous chapters. This is then followed by a general discussion including areas that might be developed as part of any future work.

2 Literature Review

2.1 Introduction

In an article published in 2011, Jay Hodgson wrote that lateral dynamics processing in current popular music practice is “as common as tapping and power chords once were in heavy metal” (Hodgson 2011). He added that hip-hop producers often used sidechain compression to convert ambient pads into rhythmic material playing the upbeat of the song to support the groove. This technique is also commonly used in electronic dance music (EDM).

However, Hodgson commented that very little attention has been paid to the technical and musical aspect of signal processing in the research on popular music practice and calls for a research program to elucidate recording practice as musical practice. Therefore, this thesis aims to shed light on the technical aspects of signal processing, particularly dynamic range compression, and its effect on audio signals, notably the timbre, when applied in a novel manner as provided by the system developed for this research.

Timbre is dependent on several attributes: the frequency content, spectral profile and temporal envelope of a sound (Houtsma 1997). One approach that is commonly used to analyse all these features is to plot the spectral profile of a sound onto a graph, also known as a spectrogram. Using a spectrogram to analyse timbre (particularly for amplitude undulations) is not a novel idea and it is an approach used in much of the literature described here. For example, one study used it to analyse the rhythmic aspects of groove by illustrating how the choice of sound can affect the listener’s perception of rhythmic variations (Danielsen 2006).

Another study used the spectrogram to compare the sound of the bass drum from 70s disco to EDM produced in the 1980s and 1990s and speculated that the success of a dance track might be dependent on the timbre of its dominant element (Zeiner-Henriksen 2006).

However, most of the published resources relating to the spectral analysis of timbre utilise short-term Fourier transforms (STFTs), including both the texts cited above (Zeiner-Henriksen 2012) (Lacasse 2012) (Solberg and Jensenius 2017) (Solberg and Jensenius 2016). The disadvantage of using this technique is that its resolution does not correspond with musical scales which have a geometric spacing (Brown 1991). An STFT uses a constant frequency difference and a constant resolution to separate the frequency content of a signal.

The process involves dividing a long-duration signal into equal length windows and equally spaced frequency bins and separately calculating the Fourier transform of each segment. Because the length of the windows is equal and the spacing of frequency bins is constant, this process produces detailed resolution for the high-frequency bands. However, for the low-frequency bands the resulting resolution is low.

This is not ideal for use in a musical context because it uses values collected using linear methods to plot data in a logarithmic plot.[2] Too little information is provided in the low-frequencies, while in the high-frequencies, the frequency resolution is too high (and the temporal resolution, therefore, is too low). The details of the resulting data do not correspond to the logarithmically spaced musical frequencies. Therefore, the next section of this thesis proposes a novel method that is more efficient as a tool to analyse timbral discrepancies. Ultimately, this method can be used to facilitate analysing the data collected from the novel mixing system developed in this research.

2.2 Method of Analysis

As mentioned in the previous section, a suitable method to visualise audio signals needs to be established to effectively analyse the effect of sidechain compression in a mix. An amplitude envelope is not suitable for this purpose as it does not provide sufficient relevant information. Consider, for example, a mix that consists of an ambient pad sidechained to a bass drum. If the

[2] Musical sounds produce identifiable harmonic patterns in logarithmic frequency domain therefore a log domain plot is more suitable for analytical purposes.

analysis is conducted using the time domain plot, the amplitude reduction on the ambient pad is not visible in the graph. This is because the point when the pad is compressed is overlapped by the dominant (bass drum) signal. Consequently, the required information to describe the behaviour of the system accurately cannot be obtained using this method.

A time-frequency representation plot (sometimes referred to as a spectrogram) is then the best option because it creates a visual representation of all the elements in the track. As part of the analytical framework for this research, the author challenges the almost universal use of the STFT for the time-frequency analysis of audio signals in production. This section demonstrates that in many cases the constant-Q transform (CQT) produces a far better visualisation (Brown 1991). Thereafter, this thesis exclusively uses this approach for time-frequency analysis. The process of obtaining the data is similar except for a few significant differences.

The first difference is the logarithmic spacing of the frequency bins used in CQT calculations, compared to the equal spacing in STFT. This means that data collected using CQT translates well onto a logarithmic graph. The linear data collected using the commonly used approach produces a lower resolution visualisation, especially for the low-frequency content. The reason for this is that the linearly produced data using STFT will then have to be enlarged, producing reduced resolution image quality in the bass frequencies. Another difference is the transform window length that is inversely proportional to the frequency.

The logarithmically spaced frequency bin and the varying transform window length produces a constant frequency to bandwidth ratio data, or in other terms, constant-Q. The result is better resolution for low-frequency content and better temporal resolution for high-frequencies (Schörkhuber and Klapuri 2010). The plot then translates well with information collected from musical sources. Furthermore, the CQT spectrogram has characteristics akin to the human auditory system whereby for low-frequency content the spectral resolution of the auditory filter is better (Moore and Glasberg 1983) and temporal information is improved for high-frequencies (Shailer and Moore 1983).

The similarities between the characteristics of CQT and the human auditory system is concurrent with a report on the low temporal resolution for low-frequency content (Danielsen 2019). It was found that low-frequency sounds will have a later placement. This means listeners will perceive low-frequency audio to occur slightly delayed. One good example of this is the high-frequency content of the bass drum in a mix. The bass drum will sound slightly delayed or not in tempo when it is missing the ‘click’ or the attack. To overcome this issue, drummers have been known to tape a credit card to the head where the beater hits to further emphasise the impact of the beater, producing audible clicks to enhance the temporal placement of the bass drum, as well as salience of the instrument in the mix (White, Robjohns and Lockwood 2013).

2.2.1 Understanding the CQT Spectrogram

In this section, a few audio examples are used to highlight the differences between the CQT and STFT spectrograms. The first one uses an 808 bass drum followed by a sawtooth wave bass synthesiser. The third example is a comparison of the aforementioned bass synthesiser and bass drum in two different states; one with sidechain compression applied and one without. A similar comparison is given after, using an ambient pad with the bass drum. The fifth example is a comparison of two approaches of using compression: sidechain and master bus. The final section presents spectrum plots taken from excerpts of three selected music examples.

These spectrograms were created using the Music Information Retrieval (MIR) Toolbox in MATLAB (Lartillot, et al.). The frequency limits were set from 20Hz to 20kHz. For the STFT, a default frame length of 50ms was used and half overlapping. The CQT plots use half overlapping varying window size. The frequency axis of the spectrograms is logarithmic because it fits the characteristics of the human auditory system. As mentioned earlier, this translates well for the CQT plots but not for the STFTs. That is why the STFT plots have low-resolution representation.

The audio synthesisers were programmed in Logic Pro X (LPX) version 10.4.8 using native LPX plug-ins. The output was normalized in LPX during rendering but was not dithered. The sample rate of the audio examples was set to 44.1kHz. An LPX compressor called platinum digital compressor was used for these tests. The reason this compressor was chosen is that it has very low harmonic distortion with look-ahead capability (Hallum 2016). It does not impart spectral colouration during compression and allows a fast attack time to be used in the examples.

2.2.1.1 808 Bass Drum

Fig. 2.1 displays two 808 bass drum spectrograms using CQT and STFT. It was set to play 120 BPM quarter notes with the velocity set to 100. The reverb send on the track is switched off, no extra processing was applied, and the fader was set to zero. As expected, the CQT spectrogram exhibits better spectral resolution in the low-frequencies compared to the STFT. Every quarter note is clearly represented, especially the duration of the note in the low-frequencies.

The visual representation of the low-frequencies also accurately represents the characteristics of the bass drum used in this example which is transient without reverberation. Additionally, spectral content from 500Hz and above are more visible in the CQT spectrogram which highlights the attack (or the 'click') of the 808 bass. This demonstrates better temporal representation compared to the STFT spectrogram.

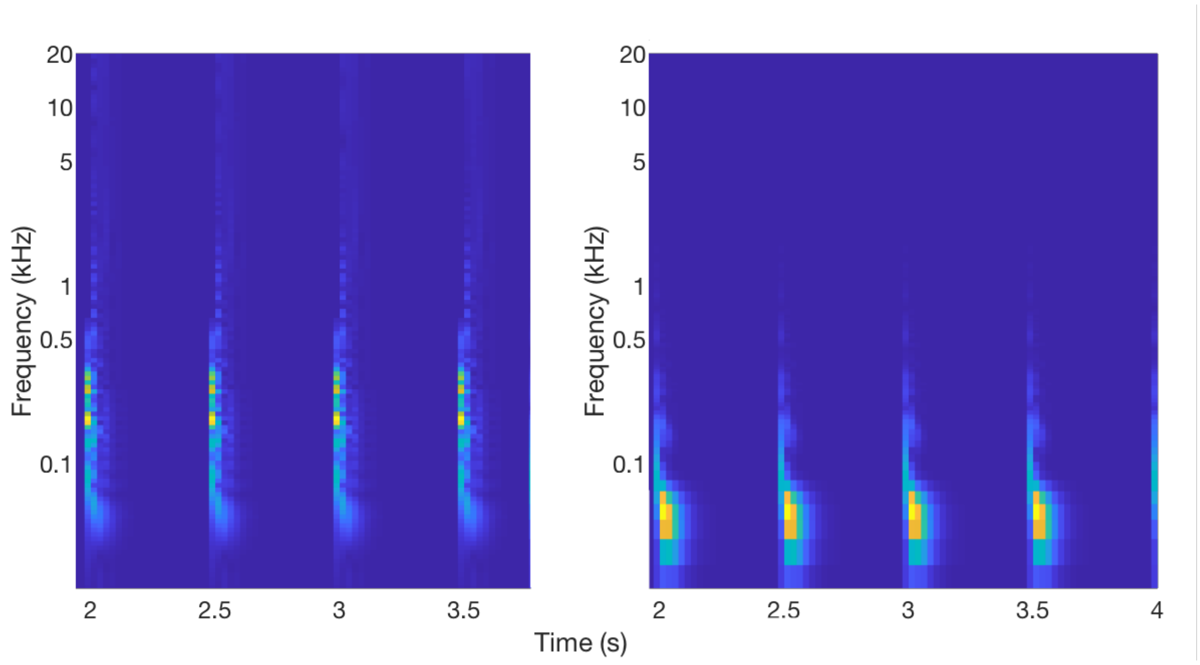


Fig. 2.1 : *808 bass drum spectrum plots using CQT (left) and STFT (right). CQT exhibits better resolution in lower-frequency and the mid to high-frequency content is visible.*

2.2.1.2 Sawtooth Bass

The sawtooth bass for this example is played at the C one octave below middle C. The synthesiser used is called 'Big Saw Bass' in LPX. The reverb send was disabled for this track and the track was converted to mono pre-fader using a plug-in called 'Gain' (Apple Inc.). The tempo was set to 120 BPM and velocity on the MIDI sequencer was set to 60. The fader was set to unity gain. The result is exhibited in Fig. 2.2. Again, as expected the low-frequency spectral resolution is better on the CQT and the temporal information from 500Hz and above are more visible. As was explained in the previous section, this is because of the longer window size of the low-frequency and smaller size on the high-frequency.

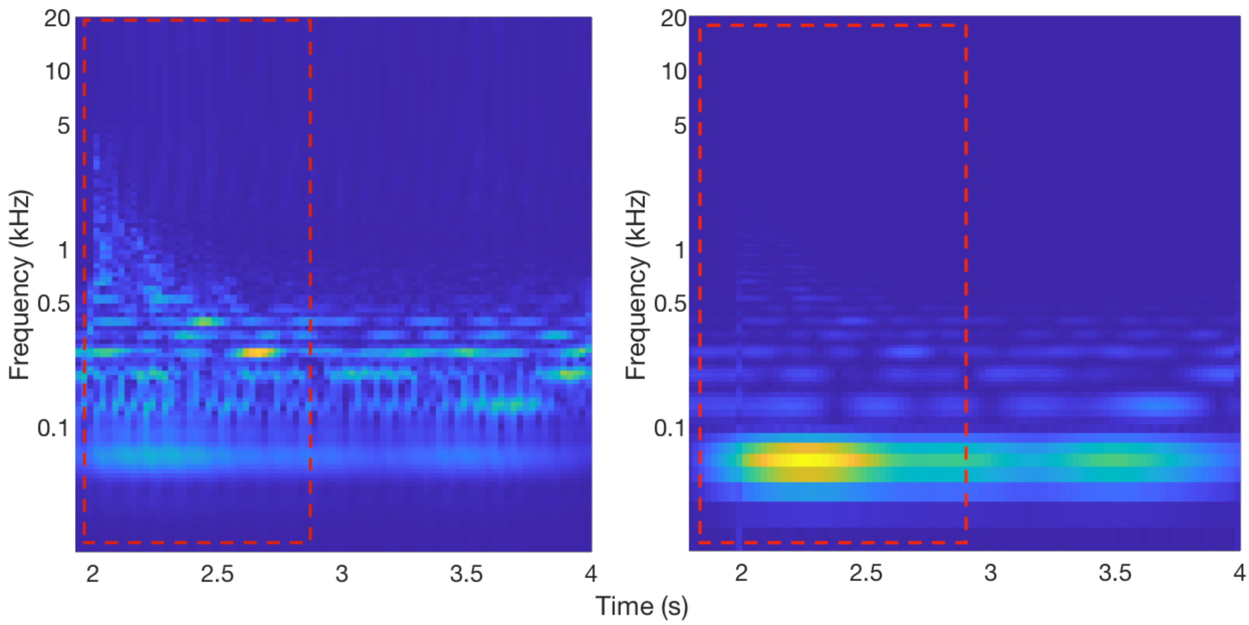


Fig. 2.2 : *Sawtooth bass spectrum plot using CQT (left) and STFT (right). The red box highlights the start of a note. CQT spectrogram exhibits better temporal and spectral resolution compared to STFT.*

For the next example, the same sample is side chained to the 808 drum track. The ratio of the compressor was set to 30:1, threshold set to -50dB$_{FS}$ and the attack and release time set to 1ms and 10ms respectively. These high values were purposely chosen to achieve heavy compression. The make-up gain was then set to 7dB to enable the sawtooth bass to have similar pre-compression amplitude peak. The result is shown in Fig. 2.3.

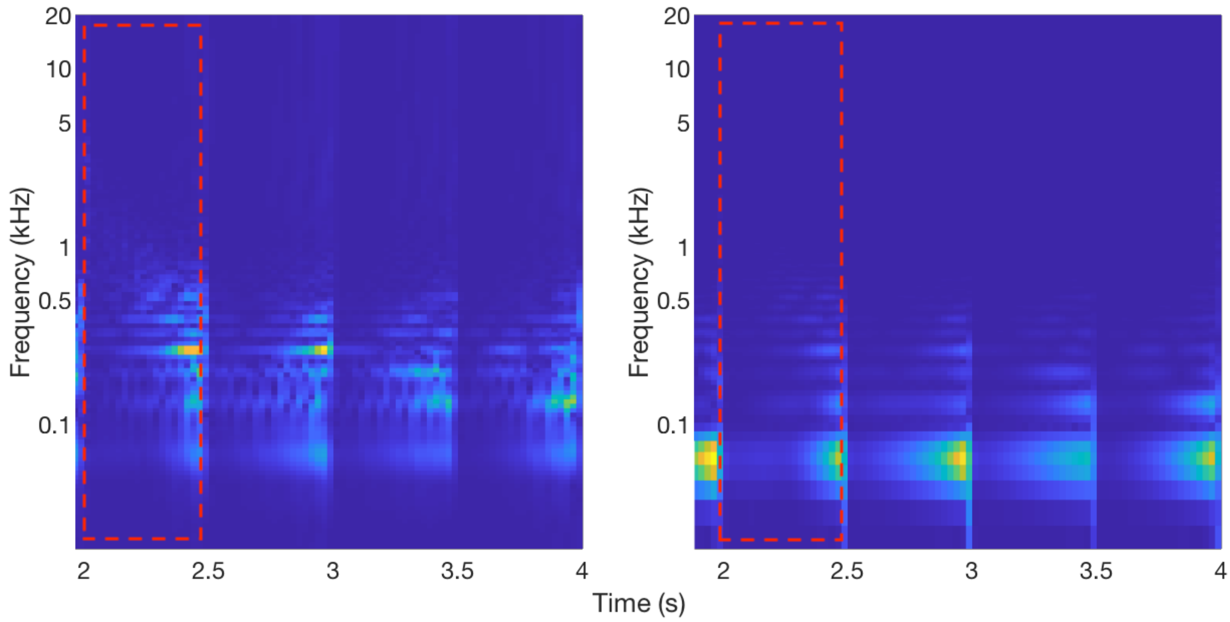


Fig. 2.3 : *Sawtooth bass sidechained to bass drum spectrum plot using CQT (left) and STFT (right). The red box is equal to the length of a quarter-note played. There is a brief silence caused by extreme compression triggered by the bass drum.*

As can be observed in both spectrograms, there is a brief cut off caused by the extreme compression triggered by the bass drum. However, new spectral components from 1kHz and above can be seen in the CQT spectrogram, but they are not visible in the STFT. Additionally, the result from the CQT spectrogram is more detailed in both the low and high-frequency spectrum. For instance, the gradual decrease of magnitude in the mid-frequencies is apparent in the CQT but is not so noticeable in the STFT spectrogram. Furthermore, the low-frequency content in the CQT spectrogram is presented with better resolution. These results demonstrate the benefits of the CQT visualisation for the effect of sidechain compression over that typically obtained via the STFT, especially for genres that put more emphasis on low-frequency content such as EDM.

For the next example, both the sawtooth bass and the 808 bass drum is plotted on the same spectrogram. The objective is to determine which spectrogram is better at representing multiple audio samples on the same plot. Based on the information obtained from the two prior examples, the visual indication given by both resulting spectrograms is intelligible (Fig. 2.4a and Fig. 2.4b), though the CQT spectrogram provides a more useful resolution, take for instance, in the low-frequency region. There, the bass drum can be easily identified in both the CQT and STFT spectrograms.

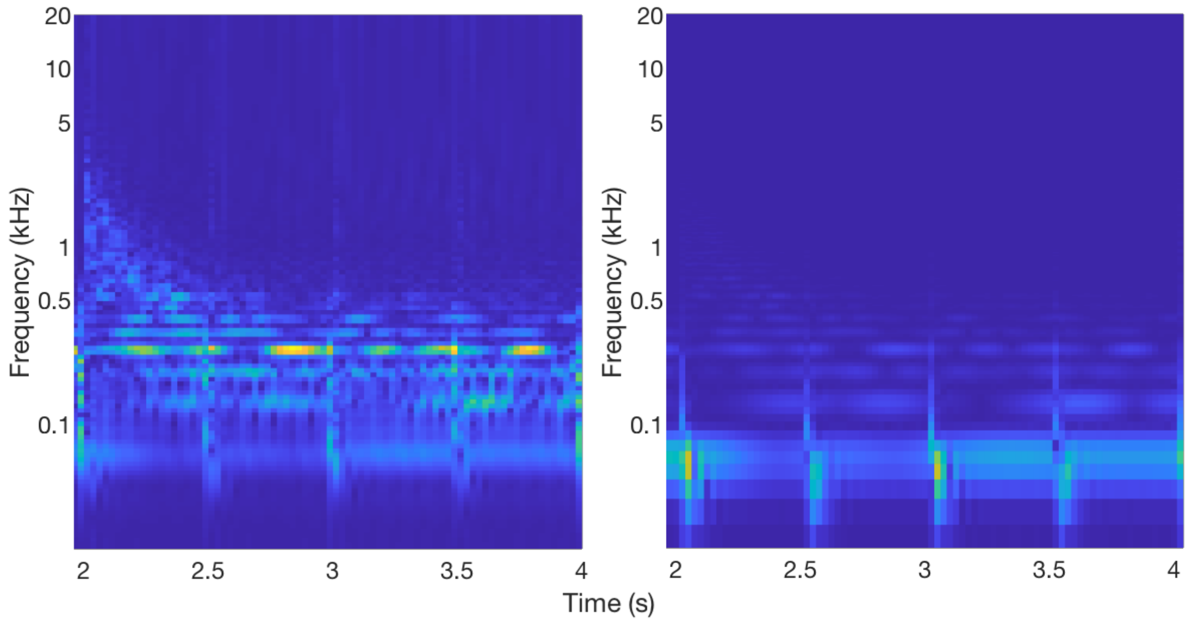


Fig. 2.4a : *Spectrograms of bass drum and sawtooth bass, not sidechained, using CQT (left) and STFT (right). Both the bass drum and sawtooth bass are identifiable in the plot.*

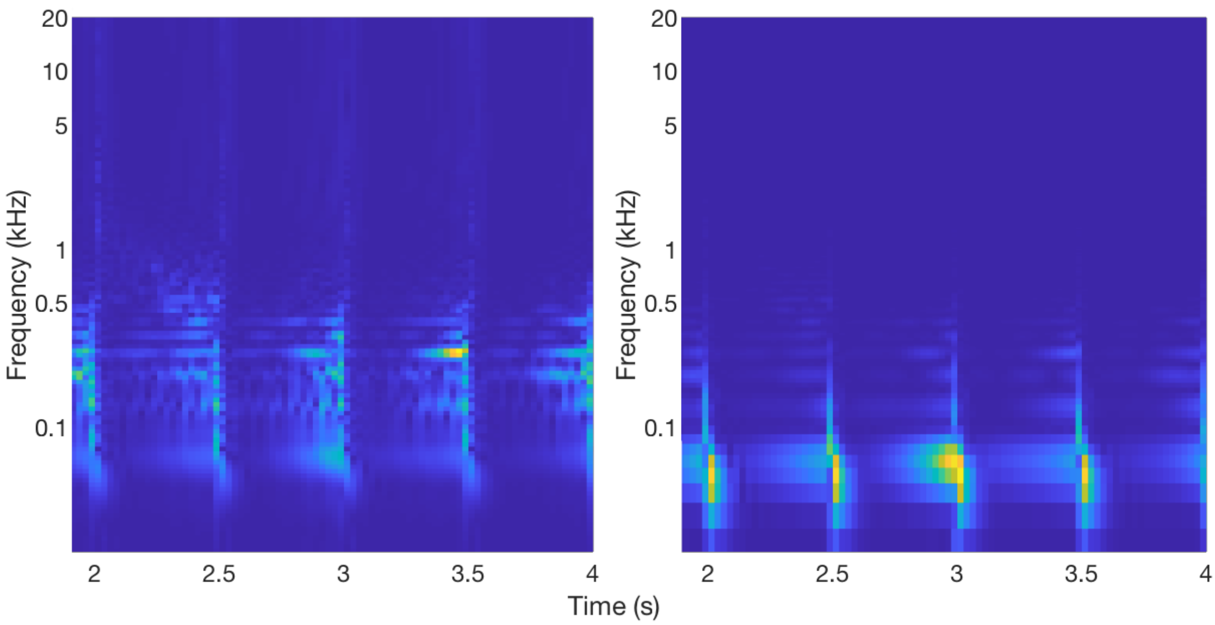


Fig. 2.4b : *Spectrograms of bass drum and sawtooth bass, sidechained, using CQT (left) and STFT (right). Both the bass drum and sawtooth bass are identifiable in the plot.*

However, the information given in the CQT spectrogram is more detailed. The bass drum is more obvious, and the length of the signal is better represented compared to the STFT spectrogram. Additionally, the temporal information for the bass drum is better represented as the high-frequency components are more visible in the CQT graph.

One obvious difference is the mid to high-frequency content that is more visible in the CQT spectrogram compared to STFT. This indicates that the CQT spectrogram produced by the MIR toolbox in MATLAB provides better spectral representation of audio samples for the purpose of analysing the content of a mix.

2.2.2 Discussion – CQT vs. STFT

The aim of this section is to determine an ideal data collection and representation method to be used in this research for analytical purposes. Considering that the CQT method provides preferable image resolution and with a resolution better suited to the signals both in low and high-frequencies, the analysis of the audio signals from here onwards will use the CQT spectrogram instead of the STFT and will be referred to only as the spectrogram. It will also be supported by other methods that fit the objective of the analysis.

2.2.2.1 Ambient Pad – Sidechain Compression

In this example, two audio samples that consist of an 808 bass drum and a synthesiser are displayed on a spectrogram (Fig. 2.5). The objective of this experiment is to compare and study the characteristics of a track that is processed using sidechain compression.

The synthesiser used is an LPX synth called Classic Ambient Pad. For this MIDI track, the velocity was set to 60, fader kept at unity gain and the note played a semibreve C major chord on middle C. No reverberation was applied to the track and it was converted to mono using the ‘Gain’ plug-in. The samples were normalised during bouncing but not dithered.

For the first experiment, no additional processing was applied. For the second audio experiment, sidechain compression was applied to the ambient pad to induce ducking. The parameters were set to 0ms attack and 10ms release with the ratio of 30:1. It was mentioned earlier that the compressor used for these experiments has a look-ahead capability, therefore it should be able to act instantaneously. There is a 10dB make-up gain applied and the threshold was set to -50dB_{FS}.

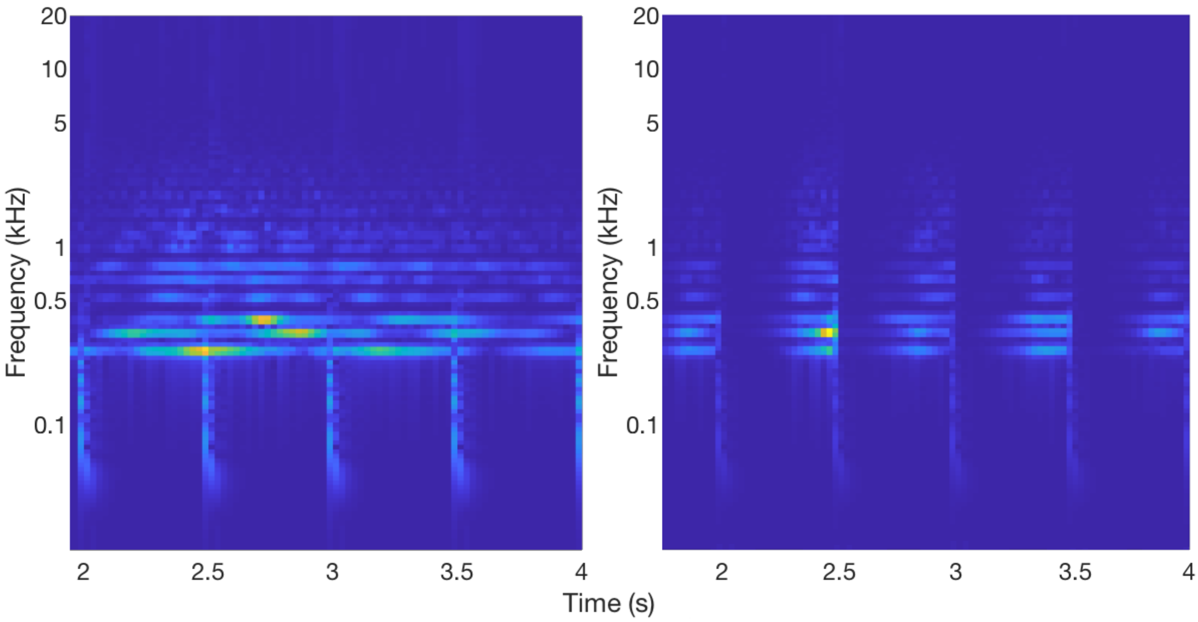


Fig. 2.5 : *Spectrogram of 808 bass drum with ambient pad; pad not sidechained to the bass drum (left) and pad sidechained to the bass drum (right).*

The spectrogram of the first experiment exhibits a clear plot of the bass drum and the ambient pad. The spectral content of the bass drum can easily be differentiated from the ambient pad even though there is significant overlap in the low-mid frequency region. The resulting plot of the second experiment, however, exhibits extreme pumping triggered by the bass drum. The brief silence after the occurrence of the bass drum may affect the temporal perception of the ambient pad.

2.2.2.2 Ambient Pad – Master Bus Compression

For this example, another way of pumping the track commonly used in EDM will be examined. The method is known as master bus compression. The difference is that in the previous experiment one track is compressed by another track through sidechain compression. This approach uses compression on the master (or summing) bus. The level of the 808 bass drum and the ambient pad used in the previous experiment is first set to unity gain. The compressor is then applied on the master bus and set to two different parameters.

The first of the examples as seen below emulates the settings of the previous experiment. This is to discover whether there are any differences between using compression in a sidechain architecture and self-compression (i.e., on the master bus). For the second example, the attack and release were set to a slower time of 100ms and 500ms respectively. The threshold and ratio are the same with the previous experiment and 7dB make-up gain was applied. The results can be seen in Fig. 2.6.

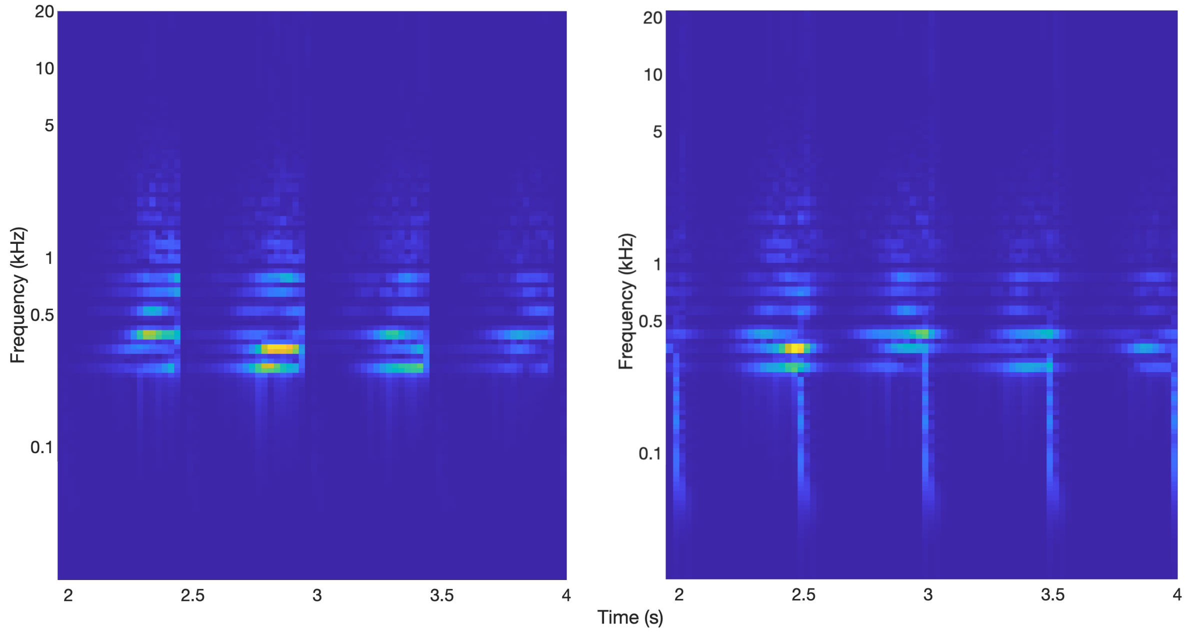


Fig. 2.6 : *Spectrogram of master bus compression with a quick attack and release time of 0ms and 10ms (right) and a slow attack and release time of 100ms and 500ms (left).*

When set to a very fast attack and release time, the track is compressed almost simultaneously as the bass drum hits. Even the distortion caused by compression is attenuated, leaving almost no spectral content in the high-frequencies. This is similar to the result produced by the sidechain architecture. The only difference is the non-existent bass drum in the plot. The spectrogram of the second example exhibits expected behaviours. The pad was present for a short duration before it was compressed.

2.2.2.3 Complex Samples – Sidechain Compression

For this example, spectrograms were generated from a mix of different types of signals that consist of 808 bass, ambient pad and a lead. The lead is a native LPX synth called Filtered Saw Lead playing a C chord, 2 octaves above middle C. An arpeggiator was applied to the lead track and summed to mono. The rate on the arpeggiator was set to a sixteenth note to shorten the duration of each note. There was no reverberation or additional processing. The faders for all the tracks were set to unity gain.

The parameters for the compressor on both tracks were set to -50dB_{FS} threshold, 0ms attack and 10ms release with the ratio of 30:1. There is a 7dB make-up gain applied to the ambient pad but none for the lead. The objective for this experiment is to observe any difference in behaviour when the sidechain system is expanded to more than two tracks and when the system process different types of signals (e.g., long sustained chord and transient notes).

For the first example, the 808 bass drum was used to compress the pad and the lead, producing interesting results. The compression did not react simultaneously for both tracks even though it was set to the same attack and release setting. This was indicated on the gain reduction meter on top of the digital channel strip in LPX (Apple Inc.). Compression constantly occurred a fraction of a second earlier on one track compared to the other. When the sequence of the tracks was set to bass drum on Track 1, lead on Track 2, and pad on Track 3; the lead is compressed a fraction of a second quicker than the pad. When the sequence of the tracks was changed to bass drum on Track 1, pad on Track 2 and followed by lead on Track 3; the pad is compressed a fraction of a second before the lead. There are three possible explanations for this peculiarity:

1) The graphics in LPX did not correctly sync the indicator on the channel strip.
2) The inherent characteristics of the synthesiser used. The Classic Analog Pad synthesiser algorithm included low-frequency oscillators (LFO) that uses a default value of 4.1Hz. The Filtered Saw Lead uses two different LFOs with a default value of 4.1Hz and 5.97Hz. Furthermore, LFOs for both synthesisers do not respond to the start/stop button. Not only it is not in sync with the tempo of the track, but it will also start at a different stage of the oscillation cycle every time. Therefore, compression could have been triggered at a slightly different LFO cycle.
3) This could be caused by sequence-sensitive block sample processing. It has been reported that LPX processes audio per sample block, and not per sample (Randolph 2019). Consequently, there could be latencies caused by sample block processing which is then amplified by sequence-sensitive behaviour of LPX.

To investigate, another experiment was then conducted using two sine waves; 1kHz and 2.5kHz, sidechained to the 808 bass drum using the same parameters as the previous experiment. The aim is to investigate whether sidechained signals are compressed at a different onset point. During processing, it was observed that the gain reduction meter was slightly out of sync. However, the spectrogram showed that the onset points are the same for both tracks. Therefore, the peculiarity might be caused by the latency on the graphics processing as well as the LFOs.

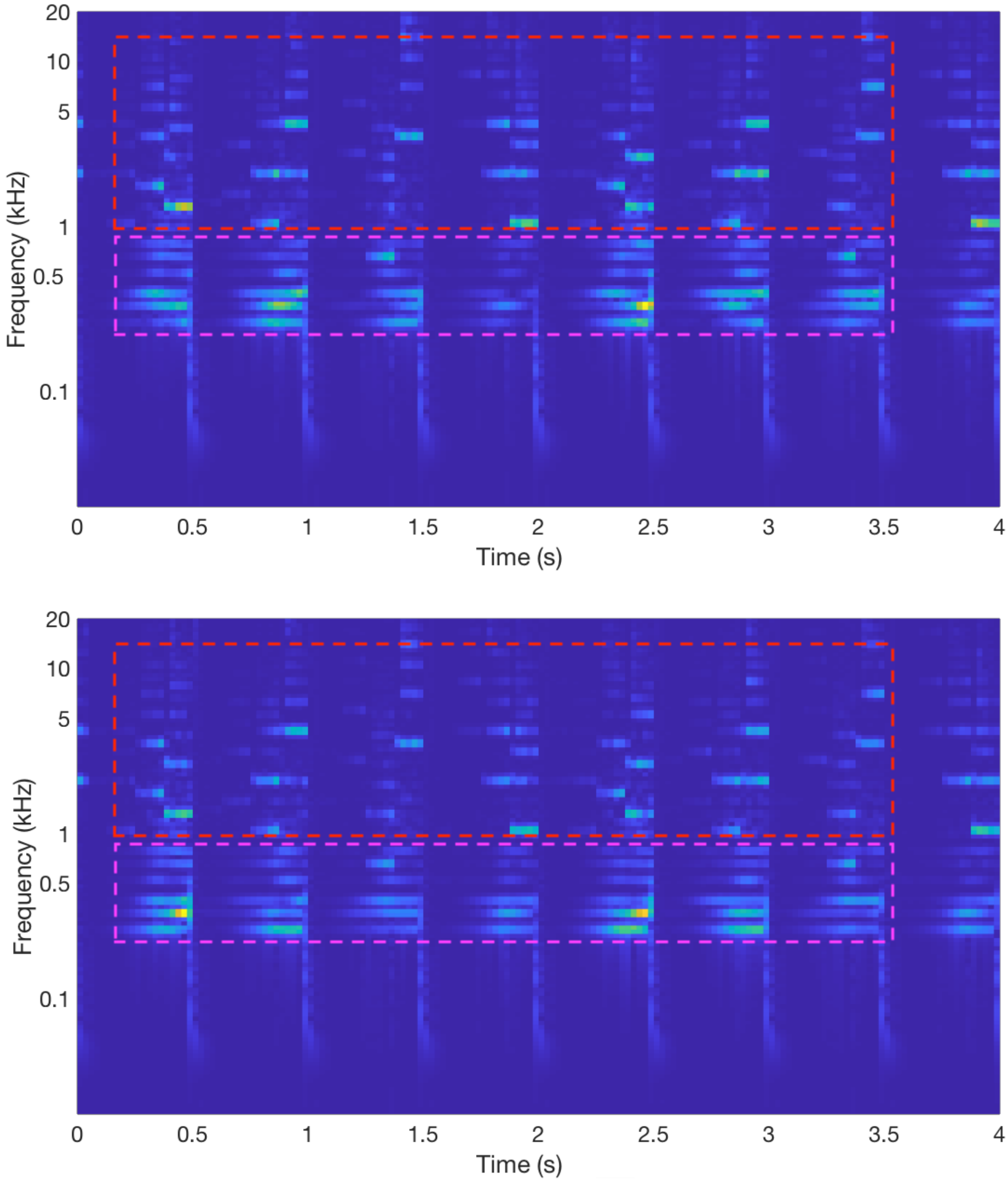


Fig. 2.7 : *Spectrogram of complex signals compressed with the same parameter values in different track sequence; bass drum, lead, and pad (top) bass drum, pad, and lead (bottom).*

Fig. 2.7 exhibits the spectrogram of the ambient pad and the lead synthesiser sidechained to the 808 bass drum. The lead is displayed in the red box while the ambient pad is in the purple box. The magnitude of the elements in the spectrogram changed slightly in both plots. One obvious example is on the ambient pad on 0.5 second mark. The same characteristics can be seen on the lead synthesiser track. This confirmed that LFOs on the synthesiser affected the audio.

2.2.2.4 Complex Samples – Master Bus Compression

For the final example, pumping was triggered using master bus compression. The platinum digital compressor was set up on the master bus, with the attack set to 100ms, release time of 500ms, threshold -50dB$_{FS}$, ratio 30:1 and no make-up gain. The objective for this experiment was to discover any unique behaviour when a master bus compression technique is used to process three different types of audio samples. The result is different from the previous example as can be seen below (Fig. 2.8).

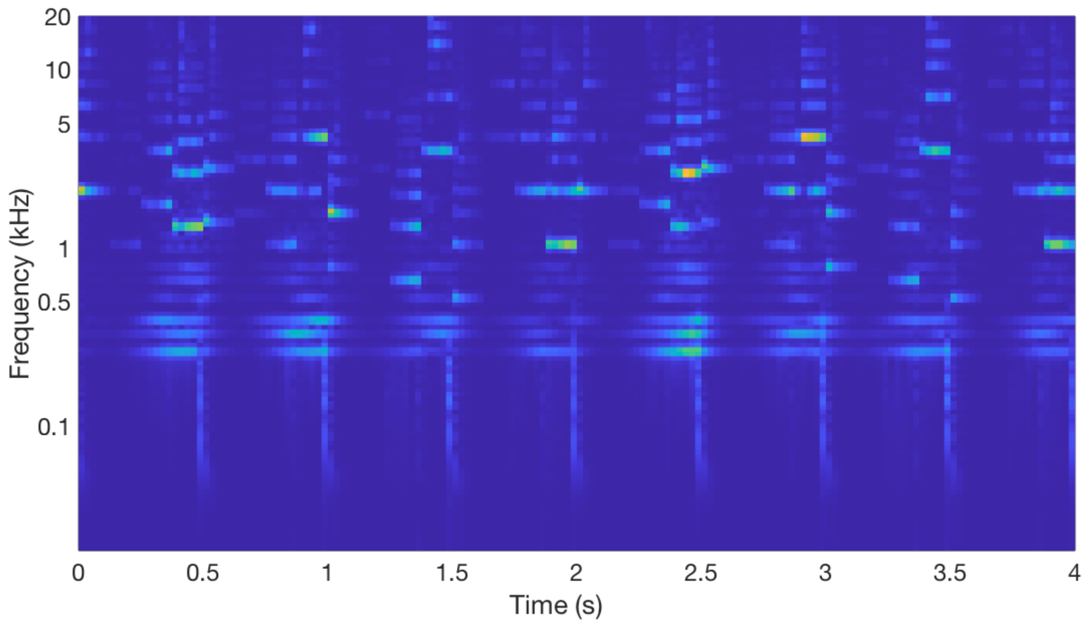


Fig. 2.8 : *Spectrogram of master bus compression triggered by bass drum. High-frequency content are still visible even after compression.*

Compression seems to affect the ambient pad more than the lead. The low-frequencies are also quieter compared to the results from the previous experiments. Furthermore, the magnitude for the high-frequencies is higher. This is because the low-frequency components are more energetic than those at high-frequencies.

2.2.3 Discussion – CQT Spectrogram

This section has explored the use of CQT spectrograms to analyse the spectral content of the different types of signals: transient, low-frequency and steady-state. It also discussed the effect of sidechain compression and master bus compression. Finally, this type of spectrogram was also used to analyse complex signals and explained a particular peculiarity that was found in one of the tests. The amount of detail revealed by the CQT spectrograms further confirms the suitability of this data presentation method for the purpose of analysing musical frequency content. In the next section, this method will be utilised to analyse the spectral content of three published musical works that features creative use of pumping.

2.2.3.1 Audio Example 1 – Music Sounds Better with You

The first example chosen for this section is Music Sounds Better with You (Stardust 1998). It was reported that this track uses the master bus compression technique (Rogerson 2019). The spectrogram is provided in Fig. 2.9, where the bass drum is indicated by a red box and gain reduction is a purple box. One feature that is apparent is the different duration of compression release. In the second purple box the gain reduction seems to happen for a longer duration compared to the other boxes. Additionally, the amount of compression also seems to be varied, indicated by the different magnitudes of the spectral content during compression.

One explanation for these characteristics is physical manipulation of the parameters during the song. Varying either the threshold or the ratio will affect the amount of compression while manipulating the release time leads to a longer rise. Tactile control of audio processors to change the characteristics of samples is not unheard of. DJs in the disco era manipulated three-way crossovers to exaggerate low-frequencies, following the beat of the bass drum (Fikentscher 2000). In present day, DJs use filters to remove low-frequencies and re-introduce them when intended during live performances (Solberg 2014). In an introspective article, the author reflected that;

> "… Live Electronic Music is a product of the belief that the body is participating once again in the music making process, that the human is having a physical effect on music, not just pressing buttons to facilitate the playback of recordings." (Vandemast-Bell 2013)

Another study summarises that risk and improvisation are indications of an authentic DJ performance and create authenticity when compared to the original material that was sampled (Rietveld 2017). Thus, by manipulating the controls a producer will have better emotional connection to the music that they produce. This is because they play a part in the composition, they are not only the enabler (or colloquially known as a 'button pusher'). The result will also be unique compared to the source of the audio samples that they have used and will introduce variation into the material. Thus, it might be the case that Stardust manually adjusted the parameters of the Alesis 3630 applied to the master bus that caused varying amount of compression throughout the track.

2.2.3.2 Audio Example 2 – Tea Leaf Dancers

The second example for this section is the song released by Flying Lotus titled Tea Leaf Dancers (Lotus 2007). It was reported that the producer used the sidechain method instead of master bus compression to achieve pumping (Hodgson 2011). As can be seen in Fig. 2.10, the spectral content for this production exhibits a stark difference compared to the previous spectrogram. The gain reduction is more dramatic, indicated by the loss in magnitude highlighted by the purple box.

Pumping that occurs in this song is constantly on the downbeat. The bass drum, however, plays a syncopated rhythm. If compression was triggered by the bass drum, compression should then be at a different rhythm and not constant downbeat. Therefore, the producer might have used another method, known as ghost sidechaining, to induce pumping for this song.

This is achieved by sidechaining the whole mix down or master bus of the tracks to an inaudible trigger track that plays constant quavers. The threshold and the ratio of the compressor on the mix down (or master bus) is then set to an extreme value. The attack is set to a value which allows the bass drum to be audible on the downbeat for a short duration before the track is pumped. Using this technique, the syncopated bass drum just before the second downbeat is not affected. The pumping is also at a constant rate, and of a constant level of gain reduction and duration. Furthermore, the possibility of any other tracks high enough in level to trigger the compression on the master bus is then eliminated. The result is different compared to the master bus compression.

The spectrogram in Fig. 2.10 represents one measure of the song. The syncopated bass drum is indicated by the smaller red box, second box from the left. Compression is not as immediate as compared to the other occurrence of bass drum. Moreover, the hi-hat in the mix does not trigger compression even though the magnitude seems to be higher than the bass drum. Therefore, it may be the case that ghost sidechain technique was used in this production whereby compression is not triggered by the bass drum in the mix but by an inaudible source.

2.2.3.3 Audio Example 3 – Can't Stop the Feeling

The final analysis is the song titled 'Can't Stop the Feeling' by Justin Timberlake (Timberlake 2016), chosen to further investigate the characteristics of ghost sidechain. For this example, ghost sidechaining is audible in the opening section of the track, particularly on the piano. The amplitude of the piano pulsates, highlighting the downbeat of the rhythm. This does not happen naturally, and it is very likely that this was processed post-recording.

The spectrogram of the short snippet of the introduction (Fig. 2.11) shows the piano track in the purple box and the hi-hat in the red box. The hi-hat plays accentuated notes on the downbeat which are clearly represented in the plot. The absence of bass drum in the mix allows for a better representation of the piano on the spectrogram because the bass drum will likely have significant spectral overlap. This is more advantageous in the sense that only the piano is represented in that spectrum which allows for an accurate analysis. Looking at the spectrogram, one could mistakenly assume that the hi-hat trigged compression on the piano. However, this is not the case because the hi-hat does not only occur on the downbeat but also the upbeat of the measure, yet no pumping can be observed on the upbeat of the piano track.

Another behaviour that can be observed is the gain reduction that occurs exactly on the downbeat. This can be seen at the onset of the hi-hat and the start of the gain reduction on the piano which looks like a straight line. This suggests the use of something metronomic to trigger the compression on the piano, with a very fast attack time and slow release on the dynamic processor. In LPX, there is a native synthesised instrument that provides a metronomic click called Klopfgeist (Apple Inc.). There are a few tutorials online that provide a guide on using this instrument to trigger compression without it being heard in the mix (vangarecord 2014) (imamusicmogul 2016). One of the reasons why it is beneficial to use Klopfgeist[3] to trigger compression is that it will not be heard in the mix, there is no reason to use any other synthesisers that might use more processing power (ibid, 1:15 – 1:30).

2.2.4 Conclusion

This section began with a brief discussion of the available methods used to analyse timbral changes caused by compression and whether they can be used to analyse the effect of compression. Two different methods of data representation for the spectrogram were then compared against one another: the SFTF and CQT. This was done by comparing the details of the spectrograms of an 808 bass drum, a sawtooth bass and the sawtooth bass sidechained to the bass drum.

The CQT spectrograms proved more effective in the analysis of dynamic variation of these signals therefore the CQT was chosen as the data gathering method to be used in this work. This section then continued with some tests to analyse and compare the spectrogram of two different processing architectures: sidechained processing and master bus processing. This was done by using an ambient pad synthesiser and an 808 bass drum. The sidechained architecture provided a different result compared to master bus processing when set to extreme parameters.

---

[3] An interesting side note, Klopfgeist in German translates to poltergeist or rapping/knocking spirit, which fits the description of this instrument for this purpose.

For the first architecture, the bass drum is not altered but the ambient pad is compressed immediately with the occurrence of the bass drum. For the second architecture, both elements are immediately compressed, including the bass drum. Changing the attack time of the compressor allows the bass drum to be present before the whole mix is compressed.

Following that was another comparison of both architectures, this time using three signals: the 808 bass drum, ambient pad synthesiser and a lead synthesiser. A peculiarity was observed in the sidechain architecture spectrogram and was investigated using another set of experiments, this time using a 1kHz and 2.5kHz sine wave. It was found that the odd behaviour might be caused by graphics issues within the DAW and also the LFO used in the synthesiser. Then, another architecture, namely the master bus, showed that compression reacted more towards low-frequency content. This is because the low-frequency components in the audio example are more energetic than those at higher frequencies.

Finally, the chapter analysed three audio snippets, ‘Music Sounds Better with You’, ‘Tea Leaf Dancers’ and ‘Can’t Stop the Feeling’. From the knowledge gathered in the earlier section, it was proposed that the producers for the first audio excerpt might have physically manipulated the compressor on the master bus. For the second example, from the analysis made in this section, this thesis proposes that a ghost sidechain compression method was used based on the constant rhythmic pumping featured in the track. For the final example, it is suggested that the producers used a metronomic short pulse to trigger compression based on the characteristics of the compression and the available literature mentioning the use of this technique for other productions.

The next section provides a brief history of compressors, particularly the four most popular types of compressors used in music production. This discussion is followed by an overview of digital compression design, particularly the design that is used in this research. This information will facilitate understanding the implementation of multi-track mutual sidechain compression in Reaper, LPX, Ableton, Pro Tools and Max/MSP.

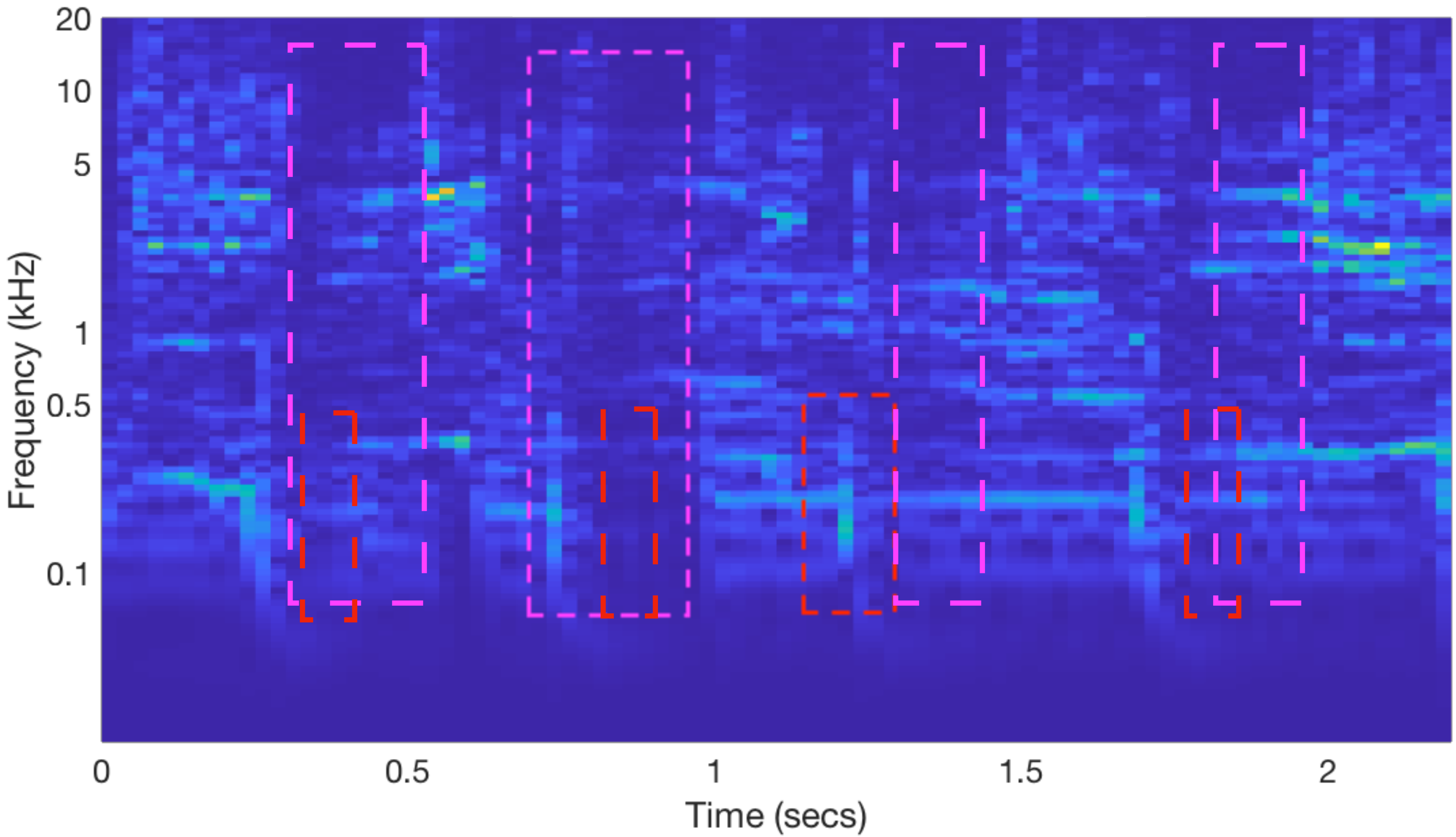


Fig. 2.9 : *Constant-Q plot of Stardust's 'Music Sounds Better with You'. The red box indicates bass drum, the purple box is the compression.*

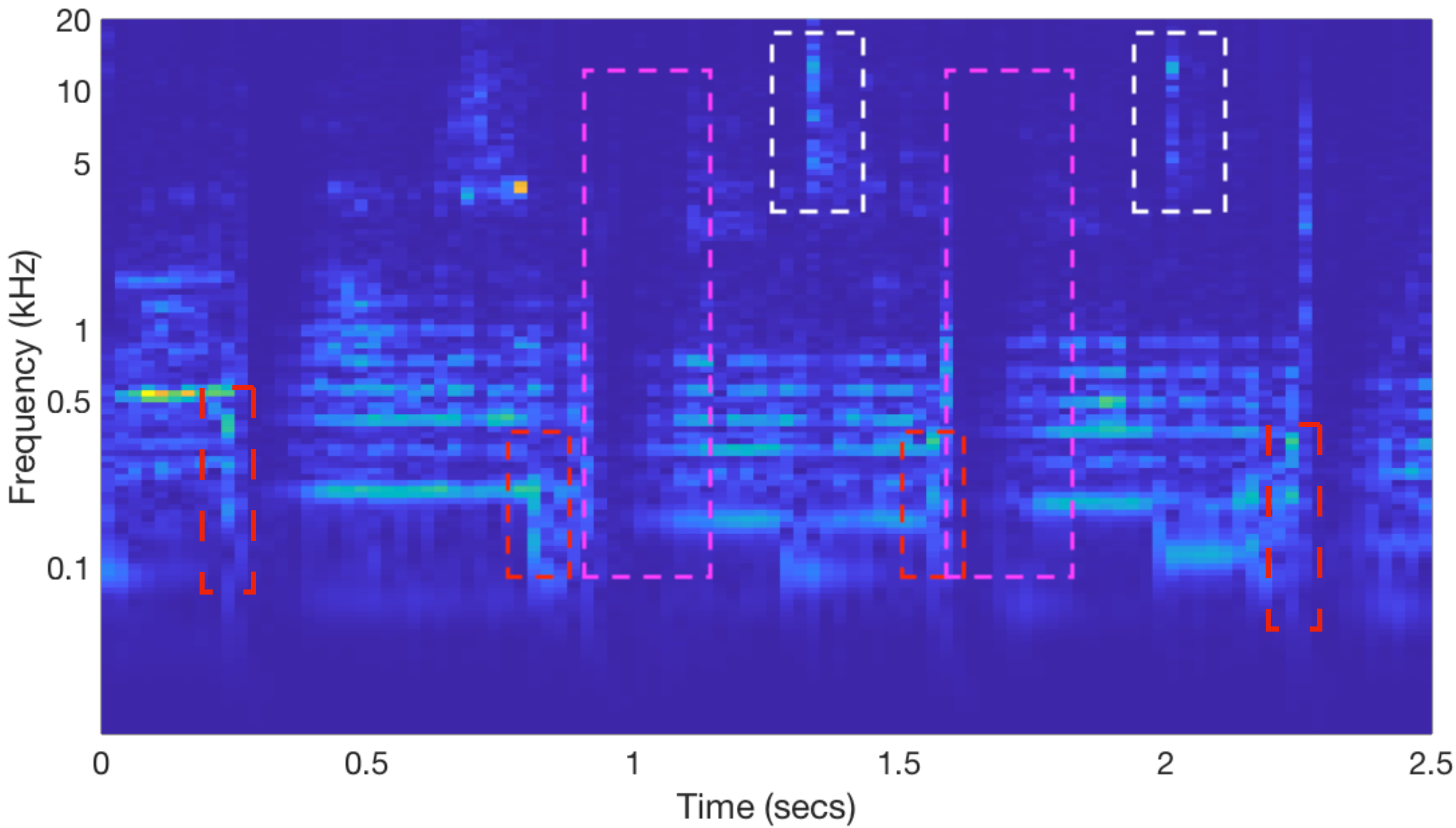


Fig. 2.10 : *Constant-Q plot of 'Tea Leaf Dancers' by Flying Lotus. The red box indicates bass drum, the purple box is the compression, and the hi-hat is the white box.*

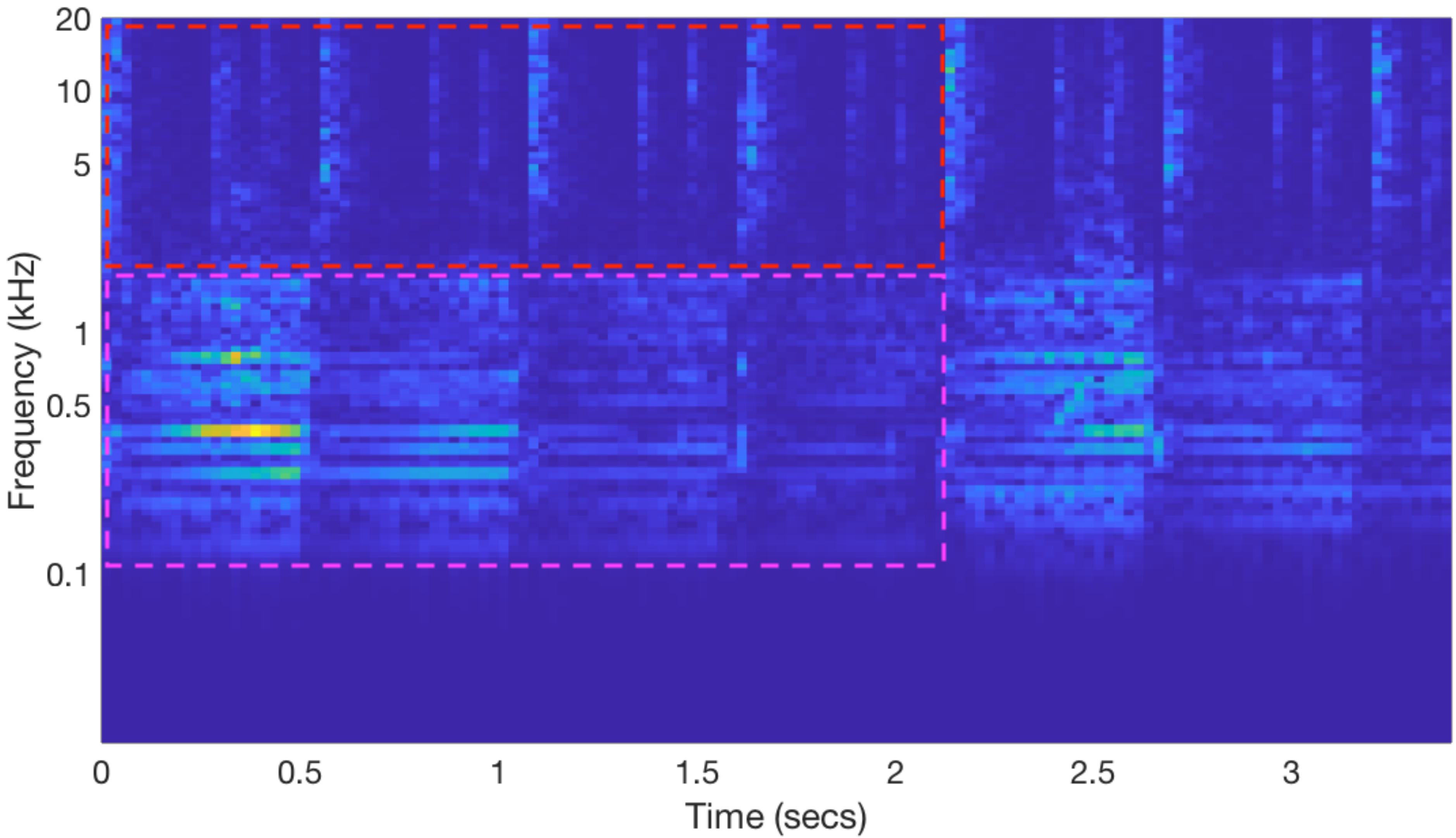


Fig. 2.11 : *Constant-Q plot of Justin Timberlake's 'Can't Stop the Feeling'. The red box is the hi-hat and the purple box is the piano.*

2.3 Automatic Level Control

One of the challenges in the early years of public radio broadcasting, specifically amplitude-modulation (AM) radio was overmodulation. It occurs when an extremely high-level signal is fed to the transmitter and overloads the modulator (Sterling and O'Dell 2010). This would often cause damage to the transmitter and the broadcast signal would be taken off the air. A sudden 3dB spike might not seem significant, but it causes a twofold increase in the power fed to the transmitter. In turn, this can cause financial losses to the broadcasting station because not only do they have to get expensive parts replaced, the broadcasting station also will have to go off the air and potentially lose listeners. Initially, this issue was dealt with using a technique called “gain riding” (Somich and Mishkind 2009). This is a simple dynamic range control process executed by a skilled engineer, adjusting the gain of a signal of an incoming signal to ensure that it does not overload.

Before compressors were invented, the ‘compressor’ in classical music broadcast was an engineer following along the musical score, behaving like a manual lookahead compressor by anticipating necessary level adjustments. However, an operator is inherently not quick enough to attenuate sudden peaks. Due to this limitation, the overall broadcast level had to leave an adequate margin of headroom to compensate for unpredictable spikes, in effect reducing the overall dynamic range of the transmission channel.

Eventually, engineers had to develop their own limiter amplifiers as a safety device, which would enable them to safely increase the loudness level of their broadcast. Other broadcasting stations soon began to implement limiter amplifiers in their signal processing chain. Noticing the demand for these devices by radio stations, manufacturers began producing the limiters commercially. Western Electric introduced the model 110A in 1937 shortly followed by RCA’s model 96A limiter (Somich and Mishkind 2009). However, even with these limiters in their transmission chain, engineers still needed to manually ride the gain on their program audio, with the limiters functioning as a safety device to prevent overmodulation.

Automated gain riding processors (now known as compressors) were then invented to complement the limiters. Used together with limiters, they raised the average RMS level of audio transmission (Somich and Mishkind 2009). The gain-control element in the early compressors was built using vacuum tubes which then changed to solid-state technology when it became more commonplace. The next section provides a brief history and description of the characteristics of compressors that are well-known in music production.

2.3.1 Compressors

> “From a technological point of view, the development of electronics thus falls in two separate epochs… the first epoch based chiefly on the principle of thermionic emission of electrons in a vacuum; the second on the conductive properties of certain materials called semiconductors.” (Collett 2003)

The advancement of electronics in the past was mostly driven by war and the defence industry. One of the inventions that came from war was the vacuum tube amplifier. Before World War I, the patent for the technologies needed to develop functioning vacuum tubes was owned by two entities, GE and Bell Laboratories.

GE’s research was focused on its potential application in voice-carrying radio systems. AT&T’s Bell Laboratories research in vacuum tube technology was focused on its application for telephone signal amplification. Eventually, the United States court ruled that the rights owned by both companies were mutually infringing. Consequently, both were denied the rights to develop their tubes (ibid, 257). This decision was overruled during World War I because of an urgent need to develop an instantaneous long-distance communication technology. As a result, research in vacuum tube technology flourished, especially in signal amplification and long-distance telecommunication. After the war ended, the demand for vacuum tubes, radio transmitters, and radio operators decreased. The ensuing surplus of vacuum tubes and available experts provided a conducive environment for public broadcasting activities and advancement in audio processing.

2.3.2 Analogue Architecture

Vacuum tubes utilise the principle of thermionic emission, whereby electrons escape from a heated metal filament electrode (cathode) and travel to another positively charged collector plate electrode (anode). This is called a diode. The first version of a vacuum tube amplifier is a triode. It is comparable to a diode, however instead of having two electrodes cased in a vacuum tube canister, there is a third electrode in between the heated filament and the plate known as a ‘grid’.

Applying a control signal to the grid would increase or decrease the number of electrons travelling from the cathode to the anode thus amplifying or attenuating the current. Other types of tubes are classified according to the number of electrodes inside the tube. For instance, a tetrode is a vacuum tube with four electrodes (anode, cathode and two grids) or a pentode which is a tube with five electrodes (anode, cathode and three grids) (Hood 1997).

One of the ways the vacuum tube is utilised in audio processing is in gain riding automation, as in the Fairchild 660. The attack and release time of the compressor depends on how quickly the electrons travel from the heated filament to the plate. Therefore, the amplification is inherently not instantaneous. Fairchild 660 has an attack time of 0.2 to 0.8 milliseconds and a release of more than 300 milliseconds (Bieger 2016).

Another war technology that was incorporated into compressor design is the optical attenuator. The T4 optical attenuator capsule was designed based on the optical sensors used for the Titan Missile Program and the Mercury Project during the Cold War (Snoman 2009). One example of how optical sensors can be used to guide missiles is by manually positioning a light source that corresponds to a two-dimensional direction on two photo-sensitive devices placed on an X-Y axis (Siekmeier 1962). The voltage generated by the photoelectric devices is used to control the missile’s servomotors and, therefore, the direction of travel of the missile.

In the case of the Teletronix LA-2A compressor, the audio signal is first amplified in the gain stage before the peak attenuation stage using the T4 capsule. Inside the capsule is an electroluminescent (EL) panel and photo-resistors. The amplitude level of the amplified audio signal is used to control the power of the EL panel and the photo-resistors control the amount of attenuation. High amplitude signals cause the EL panel to shine more brightly thereby increasing the amount of attenuation.

The attack and release time for this type of compressor are dependent on the interaction between the light source and the light-sensitive resistors. If the brightness intensity changes slowly, then the attack time will be slow. The release time, however, is affected by the EL panel's memory effect. This is because if the EL panel in the T4 has been bright for a long period, it remains illuminated for a while even after the power source has been removed (Case 2007). The average attack time for this compressor is 10 milliseconds and the release can take several minutes, depending on how long the panel has been shining (Shanks 2003).

The war effort also pushed the development of new technology to replace vacuum tubes. Tubes need a large power supply, take up space, generate heat and are heavy. Research in the field of semiconductors yielded positive results and brought forth the solid-state triode, also known as the transistor. They comprise three terminals; source, drain, and gate, which are roughly analogous to cathode, anode and grid. In valve tubes, the electrons travel from cathode to anode through a vacuum. For transistors however, the electrons travel from terminals that are sandwiched together. This enables faster transmission of electrons from source to drain. Gain control is achieved by exploiting the electric field effect. Applying a control signal to the gate will increase or decrease the number of electrons reaching drain, amplifying or attenuating current.

Transistors were introduced into compressor architecture as a replacement for vacuum tubes, notably in the UREI 1176. One of the first audio processors produced by UREI was the 176 tube compressor. The design was then revised using a field-effect transistor (FET) to replace the tubes and the processor was renamed as the 1176 (Fuston 2012). The attack and release time for this compressor is 20 microseconds to 800 microseconds for attack and 50 milliseconds to 1.1 seconds for release (ibid).

Both vacuum tube and FET compressors impart spectral colouration to the audio signal (ibid). According to available studies, this is caused by the output transformer which can be found in the circuitry (ibid) (Barbour 1999). However, FETs and tubes produce different timbral characters. Tubes create a more dominant second harmonic distortion whereas for transistors the third harmonic is more dominant (Hamm 1973).

After WWI and WWII came the Cold War. Technological advancement during this time was driven by the space race. Endeavours to be the first man on the moon by the Soviet Union and the United States laid the foundations to many modern technologies that are still in use today. Sending a man to the moon requires fast computing power with no chance for error. It also requires a smaller and lighter computer that can fit inside a space capsule. All these were made possible by further advancement in solid-state devices called integrated circuits (IC).

In 1972, the world was introduced to the first pocket-sized scientific calculator, the HP-35. It was significantly smaller compared to its predecessor which was a large desktop calculator built using discrete components (Hewlett-Packard) (Engineering and Technology History Wiki). In the same year, patent rights were granted for two necessary components for the voltage-controlled amplifier (VCA) compressor; the ‘RMS Circuit with Bipolar Logarithmic Converter’ and the ‘Multiplier Circuit’ (Ballou 2015).

Both patents are the key components of a VCA compressor, the former for RMS level detection and the latter necessary for gain control. The audio processor uses IC to calculate the RMS level of the output instead of amplification by means of the movements of electrons.

The solid-state technology used in the IC enables the VCA compressor to have a fast attack time. The API 2500, for example, has an attack time of 30 microseconds. It is not much different from

FET compressors because both use solid-state circuitry. However, VCA compressors achieve compression by means of calculating the amount of attenuation using the current-voltage exponential curve of bipolar junction transistors.

VCA compressors typically do not use transformers in their circuitry. Transformers are used to match impedances between different components in the signal path, thereby ensuring best efficiency in signal propagation. VCA compressors do not need to implement transformers in their circuitry because they incorporate operating amplifiers (op amps) in their circuit design. Op amps operate in direct current. Therefore, impedance is not an issue. Consequently, VCA compressors are known for not imparting distortion on to audio signals. Furthermore, the compression behaviour is also predictable and repeatable, due to their use of integrated circuits to calculate the amount of attenuation.

2.3.3 Digital Dynamic Range Processor Design

Now that the history of compressors and the method to analyse the data have been explored, we arrive at the compressor architecture used in this research. The digital compressor chosen as the platform for the system used in the experiments was selected because it does not impart colouration to the audio signals, it enables full customisation of the architecture and allows multi-track sidechaining to be implemented in the system. Such a system is likely to be difficult to implement in the analogue domain because of the arduous routing required. For example, an eight stereo-track system would need at least 256 patch cables.

Digital dynamic compression has been explored in depth and provides a multitude of choice for the processing architecture (Giannoulis, Massberg and Reiss 2012). The elements that can be configured are: type of level detection (RMS detector, peak detector, level corrected peak detector, and smooth peak detector); system design (feedforward, feed backward, and alternate digital feedback); and the placement of the level detector (return-to-zero linear detector, return-to-threshold linear detector, and log domain detector).

Each configuration has its own advantages and flaws. Therefore, it is entirely up to the designer to choose the type of architecture that is most suitable to achieve the desired outcome. For this research, the architecture that was used is a feedforward smooth decoupled peak detector compressor, with the detector placed in the log domain after the gain computer. Transparent processing with minimal to no distortion is needed for this research. This is because the aim is to examine the behaviour of a multi-track sidechain compression system and how it affects audio signals, without any coloration.

According to one summary:

> "Feedforward compressors are preferred since they are more stable and predictable than the feedback type ones, and high dynamic range problems do not occur with digital designs." (ibid, 406)

For a feedforward compressor, the control signal for gain reduction is the input signal sent to a sidechained gain computer (Fig. 2.12a). This could be problematic for an analogue feedforward compressor as it might not be able to accurately process signals that constantly fluctuate with a broad dynamic range.

For an analogue compressor the feedback architecture is preferable as the control signal is taken from the output (Fig. 2.12b). This means that the dynamic range has already been reduced prior to processing. The system is therefore more stable because the signal is more controlled.

However, this is not a problem with digital compressors where an arbitrarily high resolution can be chosen. Feedforward compressors are also more predictable because the gain reduction is based on the input signal and thus more stable compared to feedback processing.

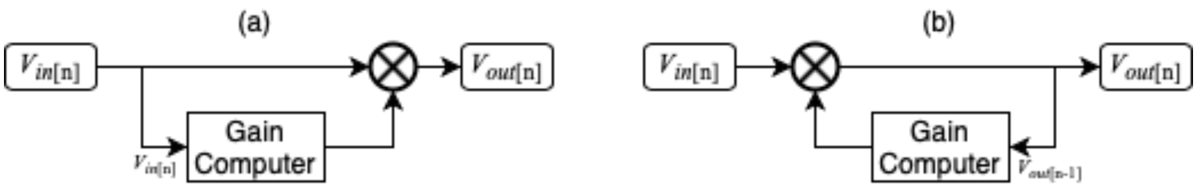


Fig. 2.12 : *Block diagrams of a feedforward compressor (a); and a feedback compressor (b).*

> “The detector is placed in the log domain after the gain computer, since this generates a smooth envelope, has no attack lag, and allows for easy implementation of variable knee width.”

The level detector is the stage where the attack and release coefficients are implemented in the algorithm. It provides a gradual change in level during and after compression or, in other words, it smooths the output of a compressor. This configuration provides the most suitable result needed for the experiment, as explained above. In addition, other placements of the level detector are mostly used for when users want to generate artefacts on purpose (ibid, 405).

> “For the compressor to have smooth performance on a wide variety of signals, with minimal artefacts and minimal modification of timbral characteristics, the smooth, decoupled peak detector should be used.”

The peak detector was chosen for this experiment because the RMS detector imposes latency as a result of calculating the average value of the recent input signal. The effect of introducing latency is not suitable for this research as it calls for transparent processing and observing the immediate effect of one signal on another.

In the analogue domain, peak detection is achieved by using capacitors. Capacitors have the ability to ‘remember’ the highest voltage value that has appeared between the plates. In the digital domain, this is analogous to the ‘max’ function where the output is the highest value of two signals. The two signals can be a combination of a number of different options depending on the chosen architecture. One example is the input level and the previous output level (Fig. 2.13).

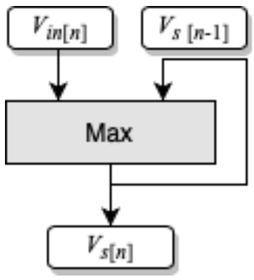


Fig. 2.13 : *Capacitor in digital domain. The maximum value for this example is taken either from the input or the feedback of the output.*

The capacitor voltage from the ‘digital capacitor’ is then calculated. This is given by the formula;

$$V_{c[n]} = V_{s[n]}(1 - \alpha)$$

Where V_c is the capacitor voltage, V_s is the supply voltage (for digital domain, it is the output of ‘Max’) and α is the time constant which describes how the voltage decays. The capacitor voltage is then used to provide gradual amount of attenuation or recovery post compression.

According to Giannoulis et. al., there are five types of level detection architectures; peak detector, decoupled peak detector, smooth decoupled peak detector, branching peak detector and smooth branching peak detector (ibid, 401-403).

For a peak detector, the output level is given by the process illustrated in Fig. 2.14.

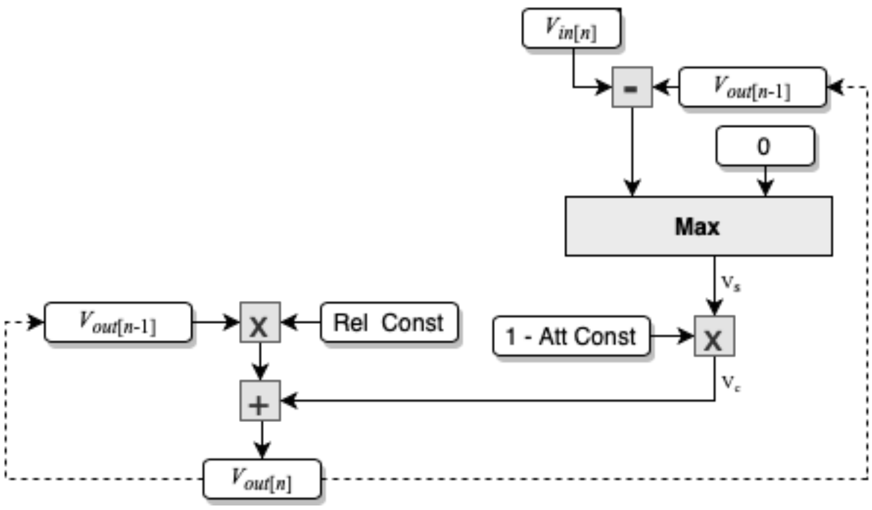


Fig. 2.14 : *Block diagram of a peak detector architecture.*

The level difference between the input and the output is first used to calculate the capacitor voltage (V_c). An input value that is lower than the output will produce a negative value. The value of V_s will then be zero, consequently the V_c will also be zero. The output will then rely solely on the release constant.

For this architecture, changes in the release time will affect the attack of the compressor because the output signal, the product of release constant and output, is fed back to the 'capacitor'. The same can be said for the attack time because the attack constant affects the V_C . The attack and release are 'coupled' because they affect one another. One way to solve this issue is to have the 'capacitor' (V_C) decoupled, as shown in Fig. 2.15. The architecture can be divided into two stages, pre-V_C and post-V_C . The attack constant is used in one stage and the release is used in another.

Using this design, the attack and release constants function independently. Therefore, any changes made to the attack and/or release time will not affect one another. If the input is bigger than zero, then the supply voltage (V_s) is dependent on the input level, meaning that the time constants operate directly on the input.

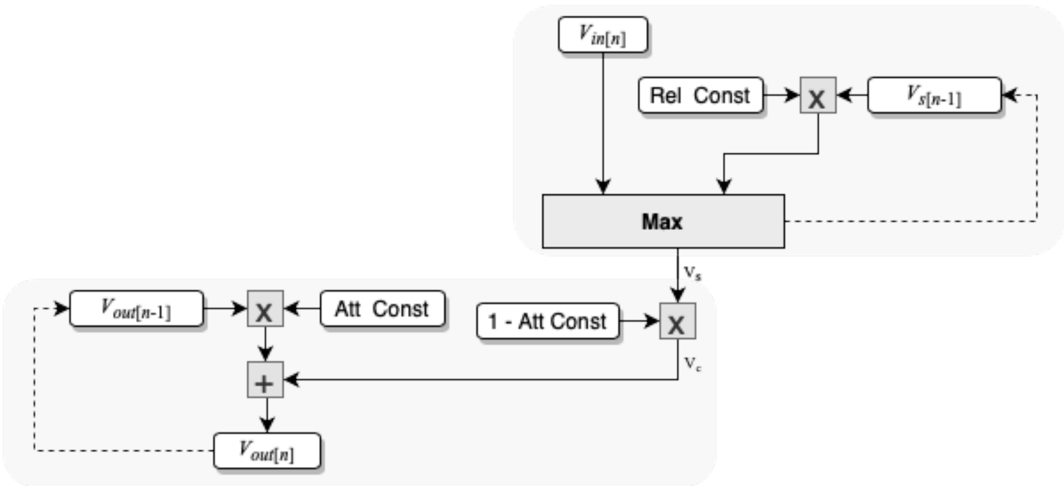


Fig. 2.15 : *Block diagram of a decoupled peak detector architecture.*

However, the full release time is applied only when the input turns to zero after a peak. If the input does not return to zero, then V_s is dependent on the value of the input thereby shortening the release time.

A smooth decoupled peak detector architecture is used when the full release time is required (Fig. 2.16). For this architecture, the input level is also taken into account for the release constant, enabling the full release time to be implemented even when the input does not return to zero after a peak. This ensures that the full release time is always applied in compression. A shorter release creates sudden stops in the envelope of the output, creating discontinuity.

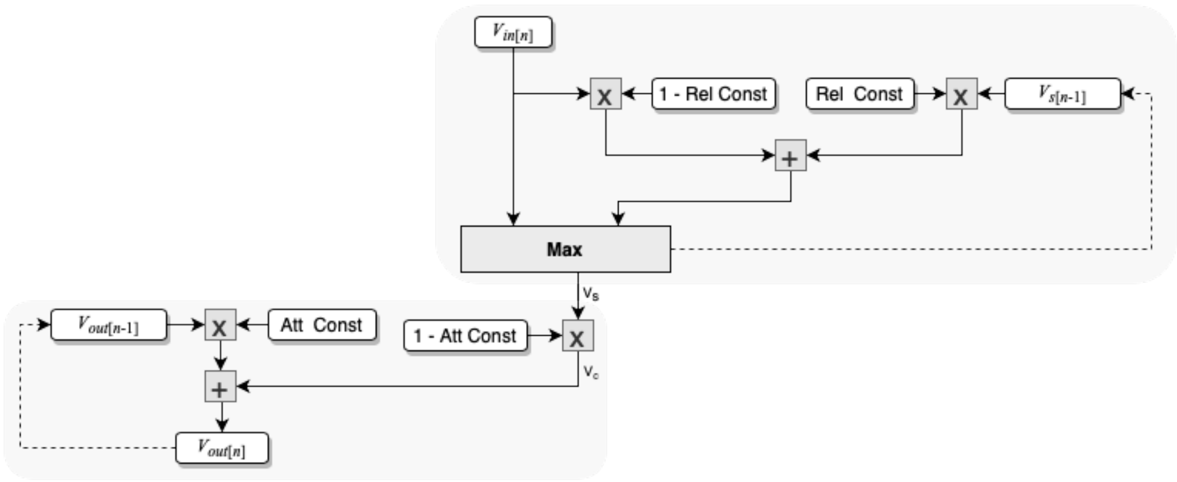


Fig. 2.16 : *Block diagram of a smooth decoupled peak detector architecture.*

A branching peak detector uses logical expressions in the architecture as shown in Fig. 2.17. The flowchart below represents a smooth branching peak detector, in which both the attack and release stage implement full release time for its process.

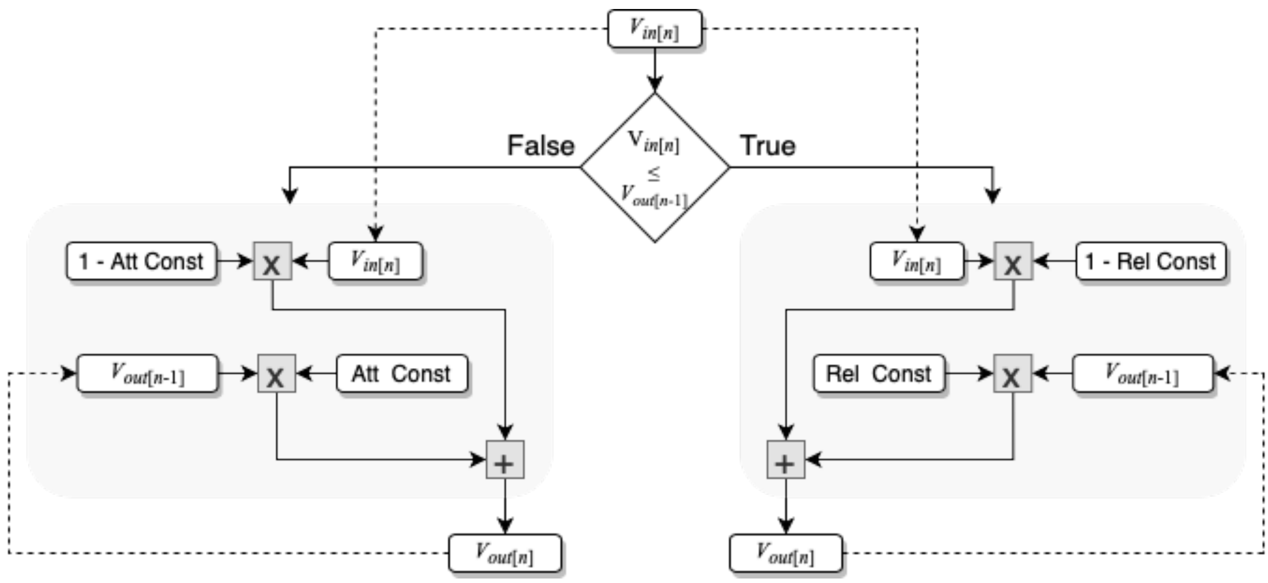


Fig. 2.17 : *Block diagram of a smooth branching peak detector architecture.*

> "Alternately, the smooth, branching peak detector could be used in order to have more detailed knowledge of the effect of the time constants, although this may yield discontinuities in the slope of the gain curve when switching between attack and release phases."

Even though the authors suggested that a smooth branching peak detector (as shown in Fig. 2.17) could be used to obtain more detail on the effect of time constants, when applied to a multi-track mutual sidechain compression the result of this might be less desirable. The reason is that the gain curve could be severely affected, causing an audio track that sounds disjointed in the mix. This is due to extreme compression from multiple sources. A sudden switch from one branch to another (attack to release) causes detachments from one state (compression) to another (post-compression). For these reasons, the feedforward smooth decoupled peak was chosen as the compression architecture for this research. Furthermore, this architecture offers transparent processing which is desirable for the system developed in this thesis.

2.3.4 Discussion – Digital Compression

This section has explored the history of compression from its inception to its current implementation. The four different approaches in analogue compression outlined in this section were driven by the advancement in analogue technology and following that is the digital processor. With these advancements came the sonic and economic implications which affect the architecture and success of an audio processor.

Sonically, valve, photoelectric and FET compressors give a distinct and often desirable aural character to the processed output, whereas VCA compressors and digital domain compressors do not impart colouration. From a financial aspect, VCA and digital compressors make dynamic range processors more affordable to obtain and maintain, while valve, photoelectric and FET compressors are costly and are usually high maintenance. These are the factors to be considered when choosing a compressor.

Nowadays, the use of dynamic range processors is not only limited to automatic gain control for corrective purposes. Since the invention of compressors, audio engineers and producers have experimented and found new ways of implementing this device in their production process, unintentionally expanding the classes of compression architecture.

Compressors belong to the category of adaptive effect processors. There are a number of ways in which adaptive audio effects may be classified (Verfaille, et al. 2011) (Reiss and Brandtsegg 2018). The author of this thesis chose to use the classification method introduced by Reiss and

Brandtsegg. This is because their classification is more detailed, taking into account the number of inputs and outputs involved in a compressor's architecture, rather than just grouping them according to the source of the control parameters. The next section explores these different classifications and anticipates a novel compression architecture called bilateral compression. This system will serve as the basis from which to understand the main focus of this research, which is a nonlinear mixing system.

2.3.4.1 Architecture 1 – Auto-Adaptive

'Auto' in this context does not refer to automatic, but 'self'. The amount of attenuation is self-regulated, based on the parameters set by the user and the analysis of the input (feedforward) or the output (feedback) of the processor. To calculate the control signal (γ_{lin}) for a stereo track, the first step is to obtain the maximum absolute value from the left and the right input channels. Then, the value is converted to the log domain.

$$V_{lin} = max(|V_{inL}, V_{inR}|) \tag{1}$$

$$V_{dB} = 20log_{10}V_{lin} \tag{2}$$

The value from (2) is then sent to the gain computer. This is where a Boolean switch is applied. The value of the threshold in decibel is given as V_T and the ratio is ρ.

$$V_{gain} = \begin{cases} V_{dB} & \text{, if } V_{dB} \leqslant l_T \\ V_T + \left(\dfrac{V_{dB} - V_T}{\rho}\right) & \text{, if } V_{dB} > V_T \end{cases} \tag{3}$$

Gain reduction (V_{red}) can therefore be calculated.

$$V_{red} = V_{dB} - V_{gain} \tag{4}$$

The value is then sent to a smoothing filter; in this case, a smooth decoupled peak detector. The attack and release time constant are incorporated in this step. This is given by;

$$\alpha = e^{-\frac{1}{\tau fs}} \tag{5}$$

where fs is the sample rate and τ is the attack or release time.

Using the value from (4) and (5):

$$V_{s[n]} = max\left(V_{red}, \alpha_{rel} V_{s[n-1]} + \left(1 - \alpha_{(att)}\right) V_{red}\right) \tag{6}$$

$$V_{c[n]} = \alpha_{att} V_{c[n-1]} + \left(1 - \alpha_{(att)}\right) V_{s[n]} \tag{7}$$

The control signal in dB ($\gamma_{(dB)}$) is given by

$$\gamma_{(dB)} = -V_{c[n]} \tag{8}$$

The final stage is to convert the control signal back to linear domain (γ_{lin}) and then to multiply the value of γ_{lin} to both the left and right channel in the track.

$$\gamma_{lin} = 10^{\left(\frac{\gamma_{dB}}{20}\right)} \tag{9}$$

$$V_{outL} = \gamma_{lin} \cdot V_{inL}$$

and (10)

$$V_{outR} = \gamma_{lin} \cdot V_{inR}$$

The compressor built for this research has a hard knee to produce overt compression. Fig. 2.18 is a flowchart that represents the process described above.

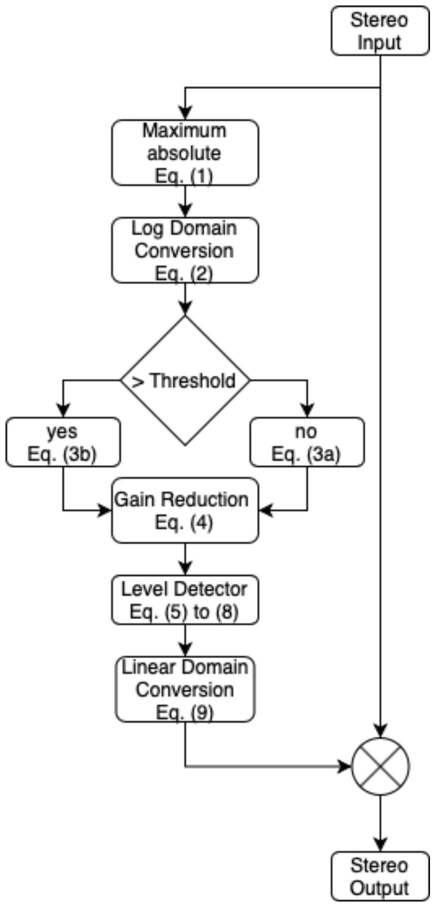


Fig. 2.18 : *Flowchart representing the architecture of a feedforward auto-adaptive dynamic range processor*.

2.3.4.2 Architecture 2 & 3 – Cross-Adaptive & External Adaptive

Another way that compression can be used in music production is using a cross-adaptive architecture that is more commonly known as sidechain compression. It was first used as a corrective tool, an example of which is the mitigation of masking by using one signal to attenuate another. This effect can be heard in Céline Dion's "The Power of Love" (Hodgson 2010). During the chorus, the vocal's reverb and delay can clearly be heard every time she pauses and is attenuated when the vocal is present in the track. This is done by using the vocal track to attenuate the reverb track through sidechain. To achieve the effect given in the example above, γ_{lin} is calculated using the output value from the 'Vocal' track using Eq. (1) to (9) and then multiplying it with the stereo 'Reverb' bus track as per Eq. (10). This is illustrated in Fig. 2.19.

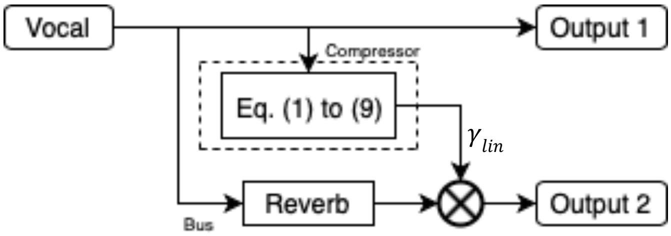


Fig. 2.19 : *Block diagram of a cross-adaptive dynamic range processor.*

A creative use of cross-adaptive compression is to create a ‘pumping’ effect, as discussed earlier. The signal flow for this method is quite similar to the one mentioned above. However, instead of attenuating a bus track (e.g., effects return), the control signal attenuates another instrument track. This can be seen in Fig. 2.20. The example given is a bass drum attenuating keyboard pad.

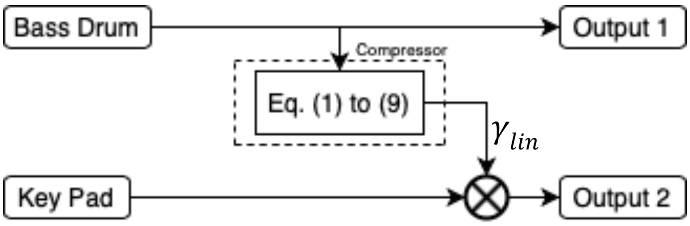


Fig. 2.20 : *Block diagram of the ‘pumping’ effect using cross-adaptive compression.*

As mentioned in an earlier chapter, external adaptive compression or ‘ghost sidechaining’, as it is commonly known, is a method whereby the element that triggers compression is not present in the mix. The process is akin to the cross-adaptive compression. The only difference is that the trigger (for the example above, bass drum) is not sent to the output. This is exemplified in Fig. 2.21.

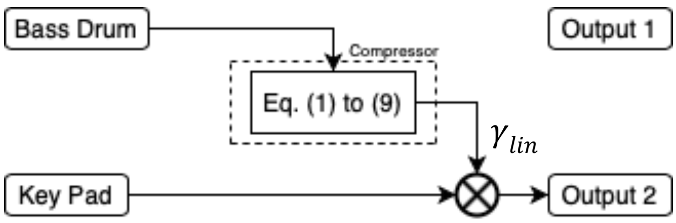


Fig. 2.21 : *Block diagram of an external adaptive dynamic range processor.*

The compression value for both cross-adaptive and external adaptive dynamic range processors can be calculated by making a few alterations to Eq. (1) to (10). The amount of compression for both the architectures are calculated based on one signal and then imposed onto another. This means the gain calculation from Eq. (1) to (9) uses the values extracted from the control (or dominant) signal which is then used to modulate the subjugated signal. Therefore;

$$V_{lin} = max(|V_{inL1}, V_{inR1}|)$$

$$\vdots$$

$$\gamma_{lin} = 10^{\left(\frac{\gamma_{dB}}{20}\right)}$$

Finally,

$$V_{outL} = \gamma_{lin} \cdot V_{inL}$$

and

$$V_{outR} = \gamma_{lin} \cdot V_{inR}$$

Fig. 2.22 shows the flowchart for cross-adaptive dynamic range processor for a two-track system. The dominant signal for this example is Stereo Input 1 (Track 1).

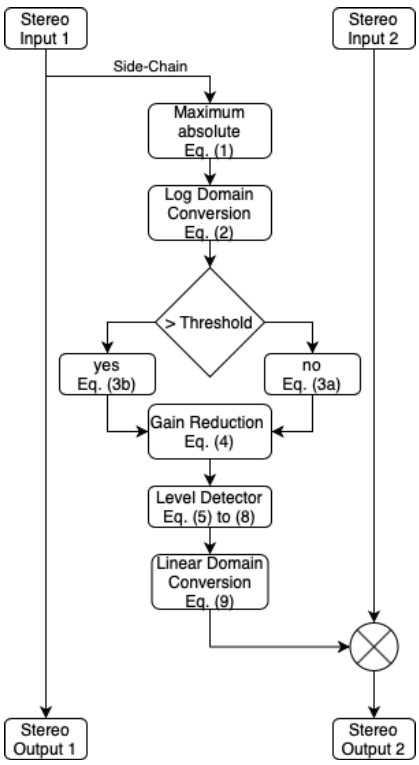


Fig. 2.22 : *Flowchart representing the architecture of a cross-adaptive dynamic range processor. The dominant signal is Track 1.*

2.3.4.3 Architecture 4 – Multi-Output Cross-Adaptive

The cross-adaptive system discussed in the previous section can be expanded to include more than one output. This is a common method used in EDM as can be heard in previously cited examples. The same algorithms are used for each track in the system. Consider a four-track system with a dominant element on Track 1. To calculate the amount of attenuation applied to Track 2 (γ_{lin2}), the values from Track 1 (V_{inL1} and V_{inR1}) are used in Eq. (1) to (9);

$$V_{lin2} = max(|V_{inL1}, V_{inR1}|)$$

$$\vdots$$

$$\gamma_{lin2} = 10^{\left(\frac{\gamma_{dB}}{20}\right)}$$

Finally,

$$V_{outL2} = \gamma_{lin2} \cdot V_{inL2}$$

and

$$V_{outR2} = \gamma_{lin2} \cdot V_{inR2}$$

The same algorithm is used to the calculate compression for Track 3, Track 4 and so forth. Track 3 for example;

$$V_{lin3} = max(|V_{inL1}, V_{inR1}|)$$

$$\vdots$$

$$\gamma_{lin3} = 10^{\left(\frac{\gamma_{dB}}{20}\right)}$$

Finally,

$$V_{outL3} = \gamma_{lin3} \cdot V_{inL3}$$

and

$$V_{outR3} = \gamma_{lin3} \cdot V_{inR3}$$

From these equations, it can be inferred that even though all the tracks in the system are compressed depending on the action of a dominant track, the amount of compression can be different depending on the parameters set by the user. Fig. 2.23 shows the block diagram of the four-track system used in this example.

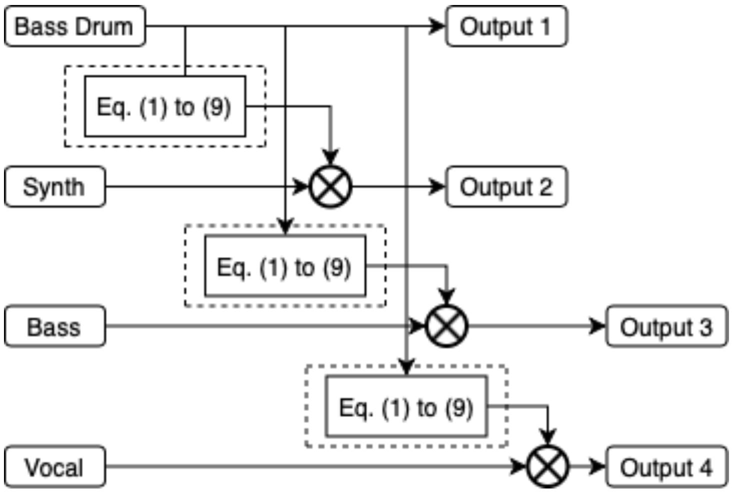


Fig. 2.23 : *Block diagram of a four-track cross-adaptive dynamic range processor.*

All these modes share the common feature that the dominant signal is not itself affected by the signal which it subjugates. Put differently, the operation is unidirectional. This can be seen in Fig. 2.15 and Fig. 2.19. In Fig. 2.15, in which the output of the dominant track (in this case 'Vocal') is not affected by the output of the subjugated track (in this case 'Reverb'). The same can be seen in Fig. 3.1 where 'Bass Drum' is not affected by 'Synth', 'Bass', and 'Vocal'. The architecture that incorporates a recursive signal flow is called bilateral cross-adaptive and this is explored in the next chapter.

3 Novel Architecture and Analysis

3.1 Introduction

So far, this thesis has covered the types of compression architecture that have been implemented in various production works, as illustrated by the examples cited with each architecture type. The next sections are the primary focus of this research. This is a detailed study of something that hitherto has not been investigated: bilateral cross-adaptive and nonlinear mixing.

3.1.1.1 Architecture 5 – Bilateral Cross-Adaptive

A bilateral cross-adaptive dynamics range processor (or referred to as bilateral compression in this thesis) is a system where two signals modulate one another, and both the processed signals can be heard. For this design, it involves a recursive feedback architecture whereby the output from one track is used as a control signal to calculate the amount of attenuation from another and vice versa. This is illustrated in Fig. 3.1.

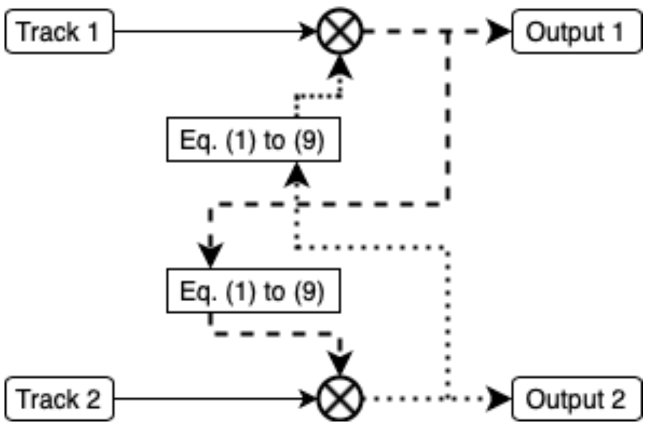


Fig. 3.1 : *Block diagram of a bilateral compression architecture. The dashed line represents the output for Track 1, the dotted line is the output for Track 2.*

Any dynamic changes from Output 1 (V_{out1}) will have an impact on Output 2 (V_{out2}) and vice versa. In other words, V_{out1} is sent to the gain computer for Track 2 and the result is used to attenuate V_{out2} . Simultaneously, V_{out2} is sent to the gain computer for Track 1 and the resulting signal is used to attenuate V_{out1}.

The difference between bilateral compression and cross-adaptive compression is that there are two control signals in the system operating bidirectionally, as opposed to unidirectionally in cross-adaptive compression.

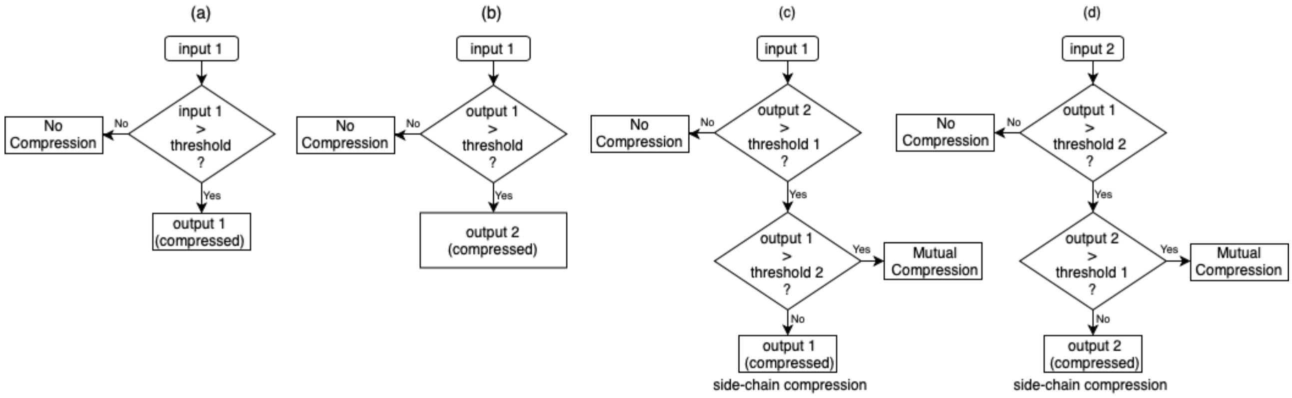


Fig. 3.2 : *Flowchart of; a) auto-adaptive system, b) cross-adaptive system, c) Track 1 in a bilateral compression system, d) Track 2 in a bilateral compression system.*

Fig. 3.2 compares the signal flow for three different compression systems. The auto-adaptive and cross-adaptive system are straightforward; if the input level exceeds the threshold, then the signal is compressed. However, the signal flow is different for bilateral compression.

Consider Track 1 in Fig. 3.2c. The amount of attenuation in Track 1 depends on how much V_{out2} exceeds Threshold 1 (V_{T1}). If the value does not exceed this, then Track 1 is not compressed. If it does, then the amount of compression is dependent on how much V_{out1} exceeds Threshold 2 (V_{T2}). In this situation, both outputs attenuate one another. In this work this condition is called mutual compression. If V_{out1} does not exceed V_{T2}, then it becomes unidirectional compression as only one condition[4] is fulfilled.

To facilitate formulating this system, consider that the attack and release are instantaneous, and the control signal is in the log domain (γ_{dB}) instead of linear (γ_{lin}). From the work cited earlier (Giannoulis, Massberg and Reiss 2012), the gain for a feedforward compressor is given by;

[4] Condition in this context means whether a signal exceeds threshold or not.

$$V_{gain} = \begin{cases} V_{dB} & \text{, if } V_{dB} \leqslant l_T \\ V_T + \left(\dfrac{V_{dB} - V_T}{\rho}\right) & \text{, if } V_{dB} > V_T \end{cases} \tag{11}$$

Whereby V_{indB} is the input level in decibels. The output (V_{outdB}) is then given by;

$$V_{outdB} = \gamma_{dB} + V_{indB} \tag{12}$$

Using these equations, we can calculate the output for conditions where only one of the thresholds is exceeded. Consider when V_{out1} exceeds V_{T2} but V_{out2} is lower than V_{T1} . In this instance, the system operates comparably to a unidirectional cross-adaptive compressor because only one track (in this case Track 2) is compressed. The same behaviour occurs for the opposite.

($V_{out2} > V_{T1}, V_{out1} \leqslant V_{T2}$).

Therefore, from Eq. (11) we can formulate gain reduction for bilateral compression for these conditions;

$$\gamma_{dB1} = \frac{1}{\rho_1 - 1} \cdot (V_{indB2} - V_{T1}) \quad \text{,if } V_{outdB2} > V_{T1} \text{ and } V_{outdB1} \leq V_{T2} \tag{13a}$$

and

$$\gamma_{dB2} = \frac{1}{\rho_2 - 1} \cdot (V_{indB1} - V_{T2}) \quad \text{,if } V_{outdB1} > V_{T2} \text{ and } V_{outdB2} \leq V_{T1} \tag{13b}$$

When both outputs exceed their corresponding thresholds, the character changes from unidirectional to mutual compression. Here, the value from both the outputs will affect the amount of gain reduction and both the tracks in the system will be compressed.

$$\gamma_{dB1} = \frac{1}{\rho_1 - 1} \cdot (V_{outdB2} - V_{T1}) \quad \text{,if } V_{outdB2} > V_{T1} \text{ and } V_{outdB1} > V_{T2} \tag{14a}$$

and

$$\gamma_{dB2} = \frac{1}{\rho_2 - 1} \cdot (V_{outdB1} - V_{T2}) \quad \text{,if } V_{outdB1} > V_{T2} \text{ and } V_{outdB2} > V_{T1} \tag{14b}$$

Eq. (13a) and (14a) can be used to calculate gain reduction for Track 1 while Eq. (13b) and (14b) for Track 2. Therefore;

$$\gamma_{dB1} = \begin{cases} 0 & \text{, if } V_{outdB2} \leqslant V_{T1} \text{ and } V_{outdB1} \leqslant V_{T2} \\ \left(\frac{1}{\rho_1 - 1}\right) \cdot (V_{indB2} - V_{T1}) & \text{, if } V_{outdB2} > V_{T1} \text{ and } V_{outdB1} \leqslant V_{T2} \\ \left(\frac{1}{\rho_1 - 1}\right) \cdot (V_{outdB2} - V_{T1}) & \text{, if } V_{outdB2} > V_{T1} \text{ and } V_{outdB1} > V_{T2} \end{cases}$$

and (15)

$$\gamma_{dB2} = \begin{cases} 0 & \text{, if } V_{outdB1} \leqslant V_{T2} \text{ and } V_{outdB2} \leqslant V_{T1} \\ \left(\frac{1}{\rho_2 - 1}\right) \cdot (V_{indB1} - V_{T2}) & \text{, if } V_{outdB1} > V_{T2} \text{ and } V_{outdB2} \leqslant V_{T1} \\ \left(\frac{1}{\rho_2 - 1}\right) \cdot (V_{outdB1} - V_{T2}) & \text{, if } V_{outdB1} > V_{T2} \text{ and } V_{outdB2} > V_{T1} \end{cases}$$

These processes happen simultaneously for the tracks in the bilateral system. The next section will explore another novel cross-adaptive architecture that is the focus for this research: a massively mutual compression system referred to for the purpose of this thesis as a nonlinear mixing system. A typical mixing approach can be seen as a linear process whereby any changes made to one track can only affect its own outcome. For example, a change in dynamics caused by reverberation in Track 1 does not affect the dynamics on Track 2 and so forth. The mixing process is therefore sequential. A multi-output cross-adaptive system, however, is only unidirectional. Therefore, changes only affect the subjugated tracks leaving the dominant signal unaltered. The mixing process is also sequential because changing the values on the subjugated tracks will not affect the others. This is different to this massively mutual compression system whereby any changes made to one track will affect another. This is because there are no dominant or subjugated tracks in this system as all tracks will have equal chances to control or be controlled. The mixing process is therefore non-sequential.

3.1.1.2 Architecture 6 – Nonlinear Mixing System

A nonlinear mixing system is an extension of a bilateral cross-adaptive architecture. It utilises more than two inputs and outputs. This is achieved by sending the output from other tracks to a summing bus and using the summation value to calculate the amount of gain reduction.

Consider a four-track nonlinear system. The outputs from Track 2, Track 3 and Track 4 will be sent to a summing bus and used to calculate gain reduction for Track 1. At the same time, outputs from Track 1, Track 3 and Track 4 will be sent to a summing bus and used to calculate gain reduction for Track 2. The same happens for all the tracks in the system (Fig. 3.3).

The signal flow for a four-track nonlinear system is shown in Fig. 3.3. The processing chain is straightforward. The output from the other tracks is sent to a summing bus before being applied to the compression equation. However, it is difficult to succinctly describe the algorithm for the process because it involves all the parameters for four different tracks. Any changes in one of the outputs will affect the amount of compression on all the other channels.

Take Track 3 for example. If the level from Track 3 is high enough to make the summing bus in Track 1 exceed its threshold, then Track 1 will be compressed. This will in return affect how much Track 1 will contribute to all the other summing busses. The mercurial nature of this system makes it difficult to represent it with a succinct summary of its behaviour.

3.1.2 Discussion – Analogue Implementation

Following the framework that has been explored above, an analogue implementation of a bilateral cross-adaptive architecture was attempted. It proved to be a difficult task. In a typical studio application, cross-adaptive compression is reasonably straightforward as it involves one signal controlling another signal in a unilateral way. Therefore, only one patch cable is needed from one output going to the other track.

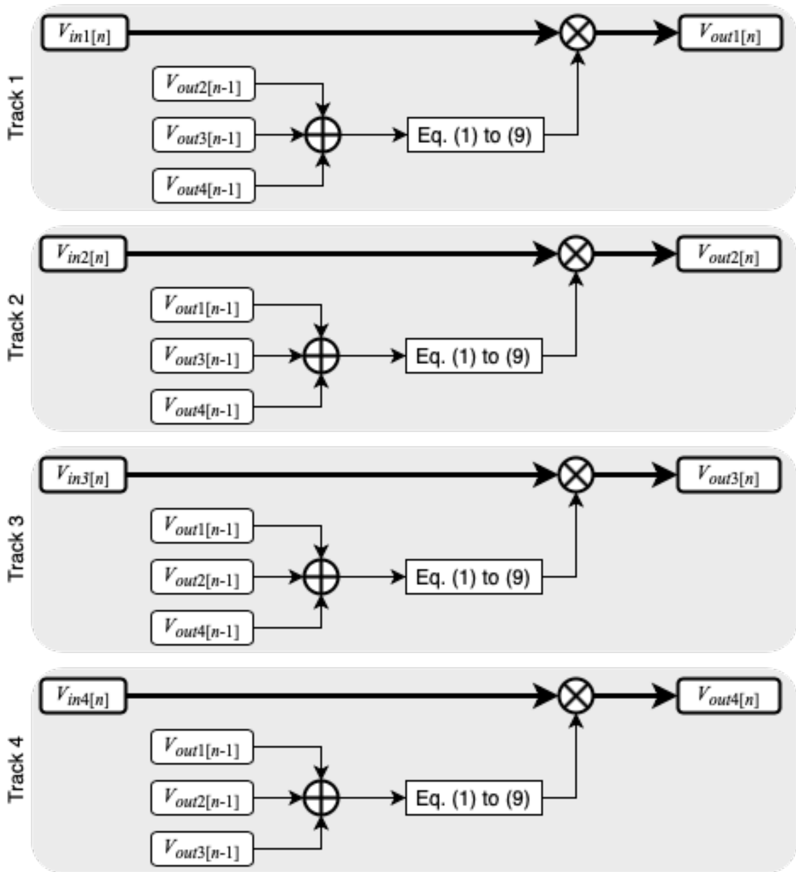


Fig. 3.3 : *Block diagram of a four-track nonlinear mixing system. The output from tracks other than itself is sent to a summing bus and used to calculate gain reduction.*

However, attempting a cross-adaptive mutual compression system is not as simple because it involves more patch cables and complicated routing. Bussing also becomes difficult because each compressor requires a unique combination of inputs from the other tracks in the system. There are also possibilities of feedback occurring in such a system when implemented in the analogue domain. The author of this thesis had attempted to emulate the signal flow using an SSL Duality Delta in an experiment to find out whether an eight-track nonlinear mixing system is possible in the analogue domain using commonly found professional mixing tools. By the time the fourth track was set up, the number of patch cables already deployed had exceeded those typically required for the entire routing set-up for a linear mixing session. This example demonstrates that it is not feasible to implement this system in the analogue domain due to the routing complexity required to implement such an architecture. Therefore, consideration of how to implement such a system with commonly available tools moves to the digital domain. This comes with its own set of challenges.

A feedback system, for a start, is usually prohibited because of the dangers of feedback. Commercially available DAWs are often used by non-audio professionals. A large number of people who purchase DAWs such as Cubase or LPX are non-experts, unlike those who would have used 24-track tape machines and a mixing desk and who more often than not would have been audio professionals. A feedback-enabled DAW might be difficult to control and produce unpredictable results which would discourage inexperienced users or beginners from purchasing it. There is also the issue of latency due to the buffer size needed for satisfactory processing efficiency. Computer-based audio workstations do not typically operate on a per sample basis, and audio tends to be processed in sample blocks. The smaller the blocks (the smallest being 32 samples), the shorter the latency. However, smaller sample blocks will cause a higher CPU load.

3.2 Implementation in a DAW

3.2.1 Deduction

Referring to Fig. 3.2, imagine a bilateral system with two steady-state sine waves placed on Track 1 and Track 2. When Track 1 exceeds its threshold but Track 2 does not, then there is no gain reduction in Track 1. This is equivalent to a cross-adaptive system following the classification in the previous chapter. The same happens when Track 2 exceeds its threshold but Track 1 does not. When both tracks exceed their threshold, mutual compression occurs. Following this logic, the compression curve of this system will have two knees with each knee corresponding to the value of the threshold set on the parameter.

However, a typical contemporary popular musical work does not only contain steady-state signals. Transient elements are common in many pop productions. To understand this, now consider a bilateral system with a steady-state sine wave in Track 1 and an amplitude modulated sine wave (henceforth referred to as AM signal) to emulate transient elements loaded in Track 2. If the output from the steady-state signal exceeds the threshold and the output from Track 2 does not, the AM signal gets constantly attenuated and the steady-state sound is not affected. This is analogous to having the fader on the AM signal constantly held to a low level.

Following this logic, the parameters of the compressors could be set in a way that will make the AM signal the dominant signal in the system and allow the steady-state sound to be modulated. The level of the threshold on the AM signal can be lowered to enable the output to exceed its value and the threshold on the other track can be set to a level high enough that it is prevented from triggering compression. As a result, the AM signal will trigger compression more often than the less dynamic signal.

The behaviour of this system, however, does not entirely depend on the threshold. It will also depend on the ratio and the release time. The track set to a higher compression ratio will be more compressed, and thus the output will be lower. It may be too low to exceed the value of the threshold, consequently affecting the behaviour and making it most likely to be the non-dominant signal. The behaviour then changes to unidirectional compression. The release time of a track will also affect the behaviour, as a longer release time will cause one track to be compressed for a longer duration than the other. This causes the level of the compressed track to stay low for an extended period which then will make the track unlikely to be dominant.

For this reason, the type of signals; whether a dynamic signal or steady-state signal, does not really matter in a nonlinear system as the behaviour of the signal can be altered by parameter adjustments. A steady-state signal can be made dynamic. A dynamic signal, however, cannot be made less dynamic because compression is not triggered by itself but by an external source. This means, the overall signal will be affected, regardless of its level.

Another characteristic that needs to be highlighted is the one indicated by the classifications cited earlier. A bilateral cross-adaptive system is defined as a system that is constructed by two signals modulating each other and a by-product of modulation is the creation of sidebands (i.e., new spectral components). Therefore, for a system to be considered as bilateral cross-adaptive, the output of the system will include sidebands which correspond to the modulated signals.

A few tests can be designed to highlight the behaviours of a nonlinear compression system. The characteristics that should be observed are as follows;

1) There should be three distinct sections in the time domain plot. Two of the sections are when the output level of one signal exceeds the threshold while the other does not. The third one is when the level of both signals exceed their respective thresholds. This is when mutual compression occurs. For a sidechain compression architecture, the result will only have two sections: compression and no compression. For a linear mix, the output will not be altered.
2) Taking into account the by-product of cross-modulation, the frequency magnitude plot for all the tracks in a bilateral cross-adaptive architecture should display sidebands that correspond to parameters of the signals used. For a sidechain system, only the non-dominant signal will display sidebands as modulation only occurs in one direction. For a linear mix, cross-modulation does not occur.
3) The interaction between the test signals should conform to the type of the signal (i.e., steady-state or transient). For the nonlinear system, when a transient signal is dominant, the steady-state signal will fluctuate. If the steady-state signal is dominant, the whole dynamic signal will be pushed down regardless of its level. If both signals exceed their threshold, then the effect on the two signals can be observed simultaneously. For the sidechain system, only the non-dominant signal is affected because the interaction is not mutual. For instance, if the steady-state signal is dominant the level of the non-dominant signal will be constantly low. However, if the dynamic signal is dominant, then the steady-state signal will oscillate. For the linear mix, no change of behaviour can be observed.

Using this information, three tests were conducted to facilitate validation of the classification of the architecture. The first test used a ramped sine wave and a steady-state sine wave to exhibit the occurrence of three different zones. The second experiment used two AM sine waves to modulate one another. This is to highlight the presence of sidebands. The third experiment used the AM sine wave to modulate a steady-state signal to observe the interactions between different types of signals.

3.2.2 Bilateral Adaptive Architecture in DAWs

Experiments were conducted to investigate whether bilateral compression architecture is supported in common DAWs. The DAWs chosen for these experiments were LPX (ver. 10.4.5) and Reaper (ver. 6.5) because of their capacity to facilitate feedback routing (Wright 2013) (ReaperBlog 2016). The signal flow in Fig. 3.1 was emulated in both workstations to conduct the tests.

The first experiment aims to observe the mutual effect of a steady-state signal to a dynamic signal in a bilateral compression system. To examine this, the test uses two 1kHz sine waves with a sample rate of 44.1kHz, set to two conditions; one ramped and one steady-state (Fig. 3.4). The parameters for the compressor in the DAW follow the values given by Table 3.1. The magnitude and the frequency of the sine wave used in this test are arbitrary.

Parameter	**Track 1**	**Track 2**
Test Signal	Steady-state (-13dB_{FS})	Ramped (From -30dB_{FS} to 0dB_{FS})
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-19dB_{FS}	-27dB_{FS}
Ratio (ρ_1 & ρ_2)	10:1	4:1

Table 3.1 : Parameters for Bilateral Compression in DAW Test 1.

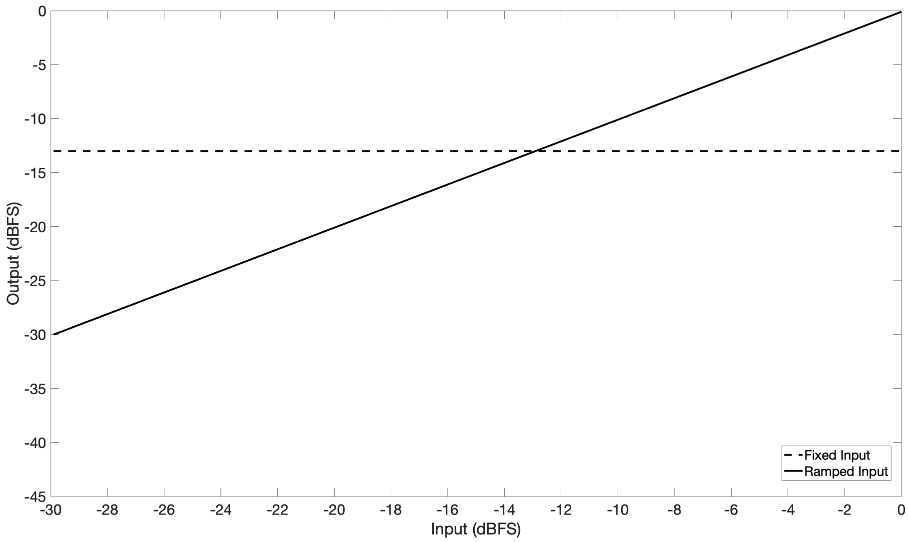


Fig. 3.4 : *Test signals for two-track bilateral compression system,* l_{out1} *(dotted) fixed to -13dB*$_{FS}$*; and* l_{out2} *(bold) ramped from -30dB*$_{FS}$ *to 0dB*$_{FS}$*.*

The second test aims to determine whether cross-modulation occurs in the DAW. To recall, one of the effects of cross-modulation of signals is the creation of sum and difference sidebands. Therefore, the sum and difference sidebands should be present in the frequency magnitude plot. This can be observed by using AM signals as test signals in a bilateral compression system.

The test signals used for this experiment were two 997Hz sine waves modulated by two proximate prime numbers: 11Hz and 13Hz sine waves (Fig. 3.5). Primes were chosen to reduce the possibility of artefacts caused by signal processing occurring at the same position on the spectrum.

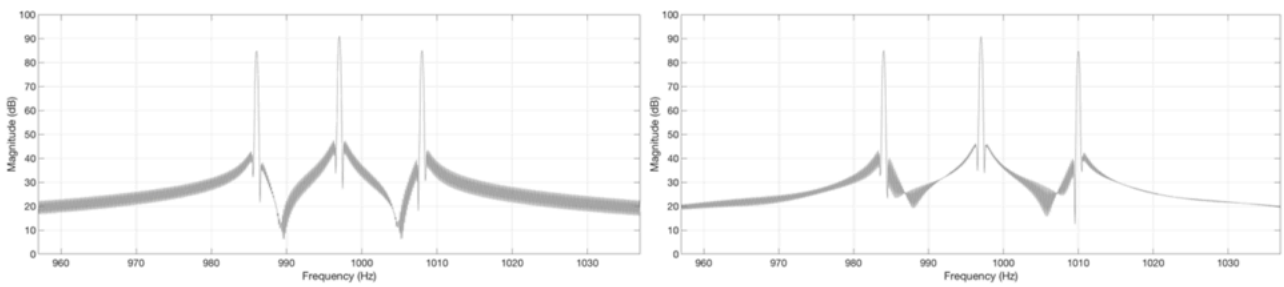


Fig. 3.5 : *Frequency magnitude plot for two test signals; Test Signal A is a 997Hz sine wave modulated by an 11Hz sine wave (left) and Test Signal B is a 997Hz sine wave modulated by a 13Hz sine wave (right).*

Subsonic modulators were chosen so that the effects of the compression can be observed as additional spectral components around the carrier frequencies. The modulator speed is set in relation to the attack and the release time so that they are relatively low compared to the time of the modulators.

The parameters for the dynamic processor were set to values listed in Table 3.2. Extreme values were used to achieve overt compression.

Parameter	Track 1	Track 2
Test Signal	997Hz x 11Hz sine wave	997Hz x 13Hz sine wave
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-30dB_{FS}	-30dB_{FS}
Ratio (ρ_1 & ρ_2)	20:1	20:1

Table 3.2 : Parameters for Bilateral Compression in DAW Test 2.

3.2.2.1 Logic Pro X

The first test was done in LPX using the test signals and the parameters mentioned previously. The compressor used in this experiment is the native LPX platinum digital compressor. This processor, which does not mimic the behaviour of any existing analogue compressors, was chosen to avoid colouration of the signal (aside from the modulation effects). The processor was set to a hard knee.

The result for this test can be seen in Fig. 3.6 (page 62). There seems to be a coaction between the two tracks in the system, indicated by both the curve on the steady-state signal and the ramped signal. The magnitude of the ramped signal is reduced from the initial level of -30dB_{FS} down to -48dB_{FS} and the steady-state was also gradually reduced to -28dB_{FS} towards the end of the test. However, the figure does not have two distinct knees and the curve does not seem to have any correlations to the parameters set on the compressor.

For the result of the second test which can be seen on Fig. 3.7 (page 62), the frequency magnitude plot from the experiment exhibits no sidebands, thus implying that cross-modulation is not a function supported in LPX. Therefore, it can be concluded that the bilateral cross-adaptive system cannot be implemented in LPX.

3.2.2.2 Reaper

In Reaper, there is an option that enables feedback routing for advanced users. However, this option comes with a caution that it might cause lower performance and loud noises. Just like the previous test, the test signals and the parameters are the same as previous experiment.

The compressor used in this test is a native Reaper compressor called ‘Major Tom’. It was chosen because the code used to build the compressor is provided in the DAW (it is implemented in the JS scripting language) and aids in understanding of how it processes gain reduction. Other than that, the system can be customised to feedforward or feedback unlike other processors that give no indication of its architecture.

The results from this experiment are shown in Fig. 3.8 and Fig. 3.9 (page 63). When the feedback routing was enabled, the system produced loud noises as soon as the track began playing. This can be seen in Fig. 3.8 where additional peaks of ≈97Hz interval can be seen. The cause of this noise is unknown to the author.

With the feedback turned off, no imprinting or cross-modulation can be seen in the frequency magnitude plot for both outputs (Fig. 3.9). These test results suggest that bilateral cross-adaptive architecture also cannot be implemented in Reaper.

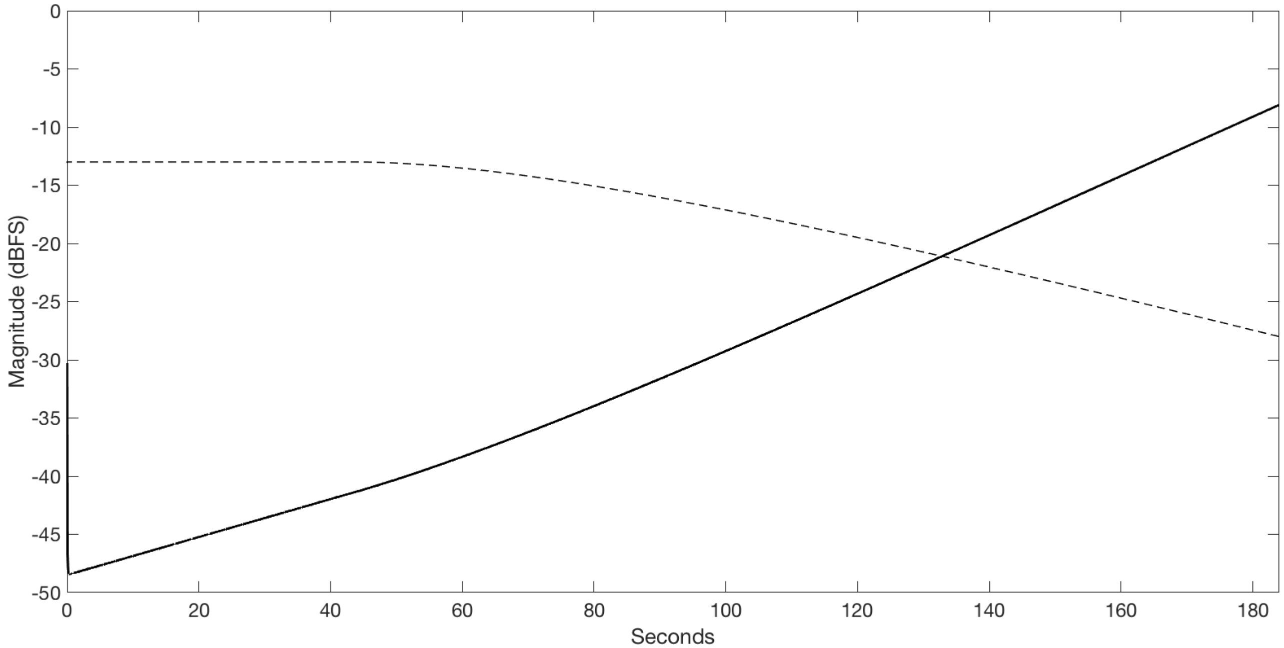


Fig. 3.6 : *Magnitude over time plot of the output from a bilateral cross-adaptive compression architecture implemented in Logic Pro X.*

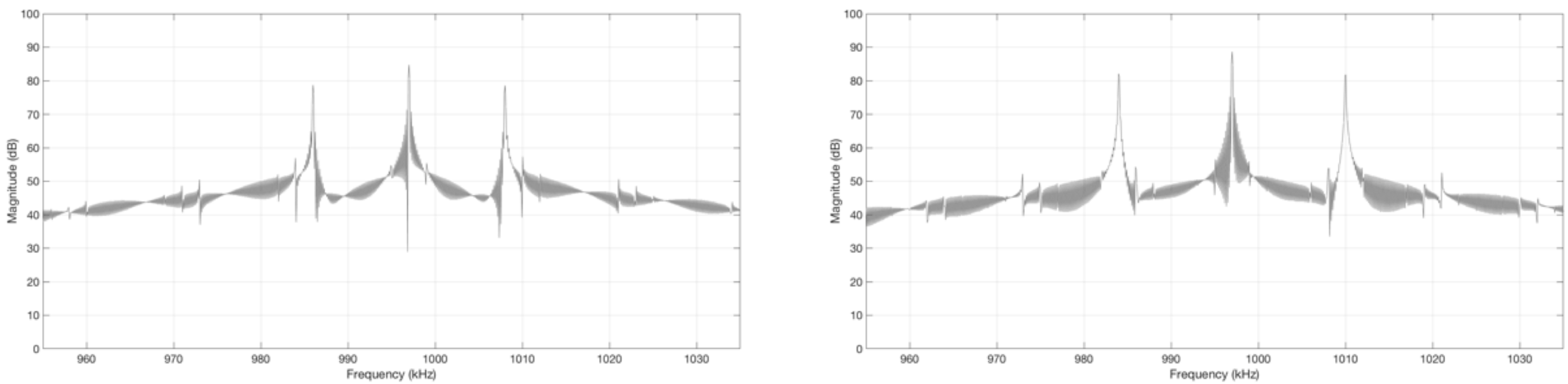


Fig. 3.7 : *Frequency magnitude plot of the bilateral cross-adaptive architecture in Logic Pro X. 11Hz modulated test signal (left) and 13Hz modulated test signal (right).*

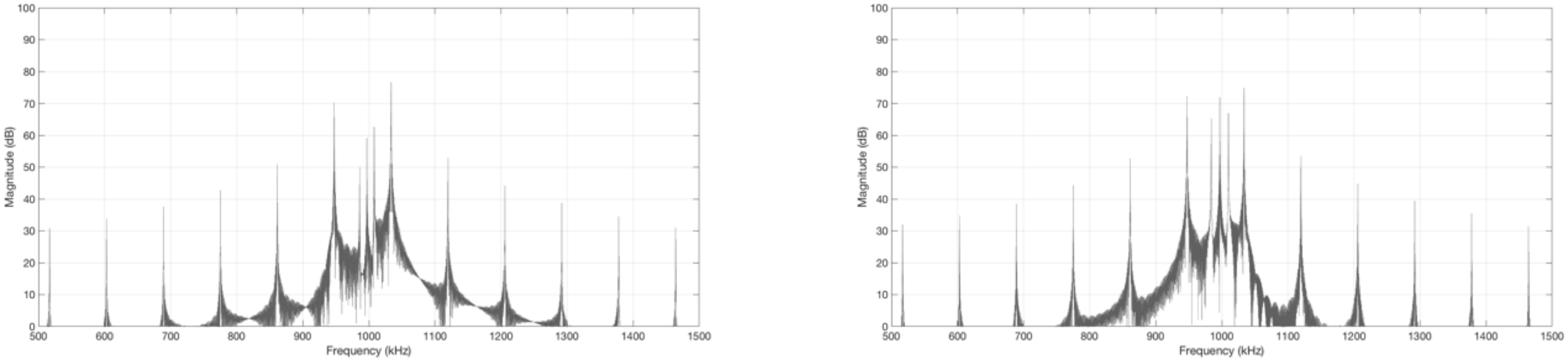


Fig. 3.8 : *Frequency magnitude plot of the bilateral cross-adaptive architecture in Reaper with feedback routing enabled, 11hz modulated test signal (left) and 13Hz modulated test signal (right).*

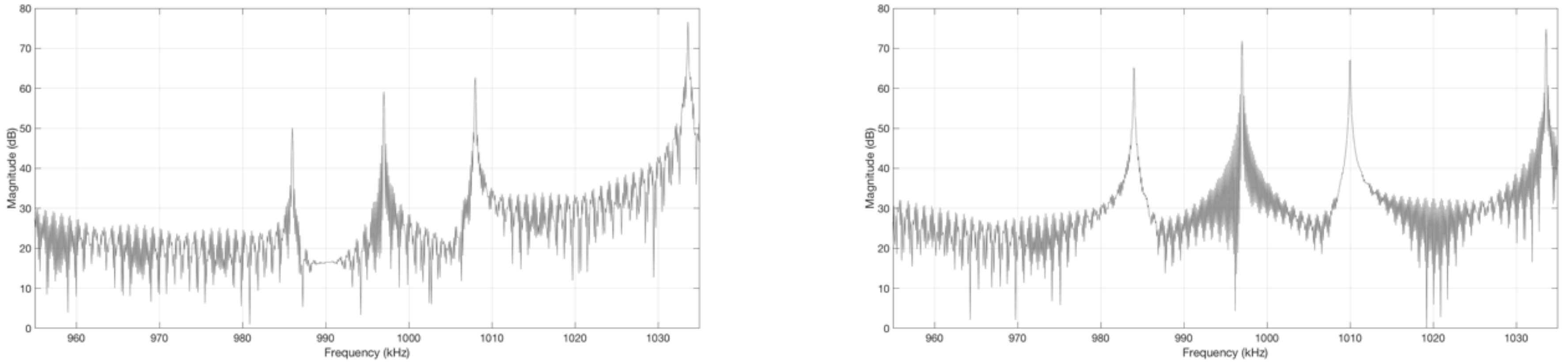


Fig. 3.9 : *Frequency magnitude plot of the bilateral cross-adaptive architecture in Reaper with feedback routing disabled, 11hz modulated test signal (left) and 13Hz modulated test signal (right).*

3.2.3 Discussion – Bilateral Compression in DAWs

The result from the experiments conducted on both DAWs proved that bilateral modulation is not possible in LPX and Reaper. The same tests were also conducted in Ableton and Pro Tools and produced similar results.

The absence of cross-modulation sidebands suggests that these DAWs process audio tracks one sample block at a time separately on each track concurrent with the report in cited earlier (Randolph 2019). This means the amount of gain reduction applied to the dominant track is calculated for the span of a window size before the value is sent to the subjugated tracks and applied to a sample window.

For self-compression, the effect of gain reduction is not obvious because latencies due to block size can be accommodated more easily where there are no complications caused by signals feeding back on themselves. This is not the case for bilateral compression system. The author of this thesis has requested either clarification or confirmation from the developers of LPX and Reaper via online forums and e-mails regarding these behaviours but is yet to receive a reply.

Therefore, to investigate the behaviour of bilateral compression system the architecture was implemented in Max/MSP. A bilateral compression system can be implemented in Max/MSP using the gen~ patching environment. This is because the gen~ environment allows per sample processing (Taylor 2018). A window size of one sample is fast enough to cover any sudden dynamic changes in Track 1, calculate the gain reduction on Track 2 using the data and send it back to Track 1. There is only a delay of two samples before a cycle is complete, which is 4.5 μs.

3.3 Implementation in Max 8

3.3.1 Introduction

In this section, a bilateral cross-adaptive architecture that has been successfully created is described. Following this, a comparison between the compression architecture with the signal flow chart is provided, followed by an exploration of a bilateral compression architecture and finally an 8-track nonlinear mixing system design and its implementation in Max 8.

3.3.2 Compression in Max 8

Fig. 3.10 is a screenshot of the inner workings of the compressor. The arrows point to the stages in which the process is represented by the flowchart. On the top right of the image, the stereo inputs from the track (boxes labelled ‘in 1’ and ‘in 2’) are simultaneously sent to sidechain processing to calculate the control signal for gain reduction, and to the output (box labelled ‘out 1’ and ‘out 2’). The boxes represent the process that will be applied to the values sent to it. The values from ‘in 1’ and ‘in 2’ are between 1 to -1 and are in the linear domain.

The first process is obtaining the absolute value of the input signals using the ‘abs’ object and the maximum value from the left or right channel is used in the calculation. This will give 100% stereo link for both channels. This method is used for the compressor algorithm in Reaper (Stillwell 2006).

The value is then converted to log domain (in dB_{FS}) before it is sent to a Boolean switch object (labelled ‘switch’). The object functions as a safety net to prevent the system from creating an error (or non-available number) when it calculates the $\log_{10}$ of 0 as the first few seconds in an audio track is usually consists of silence (or a value of zero). If the value sent is not an available number (‘isnan’ object), then the output will send the value of -200dB_{FS} instead (which is well below the noise floor of any digital/analogue conversion system). If not, it will send the level of the audio in dB_{FS}

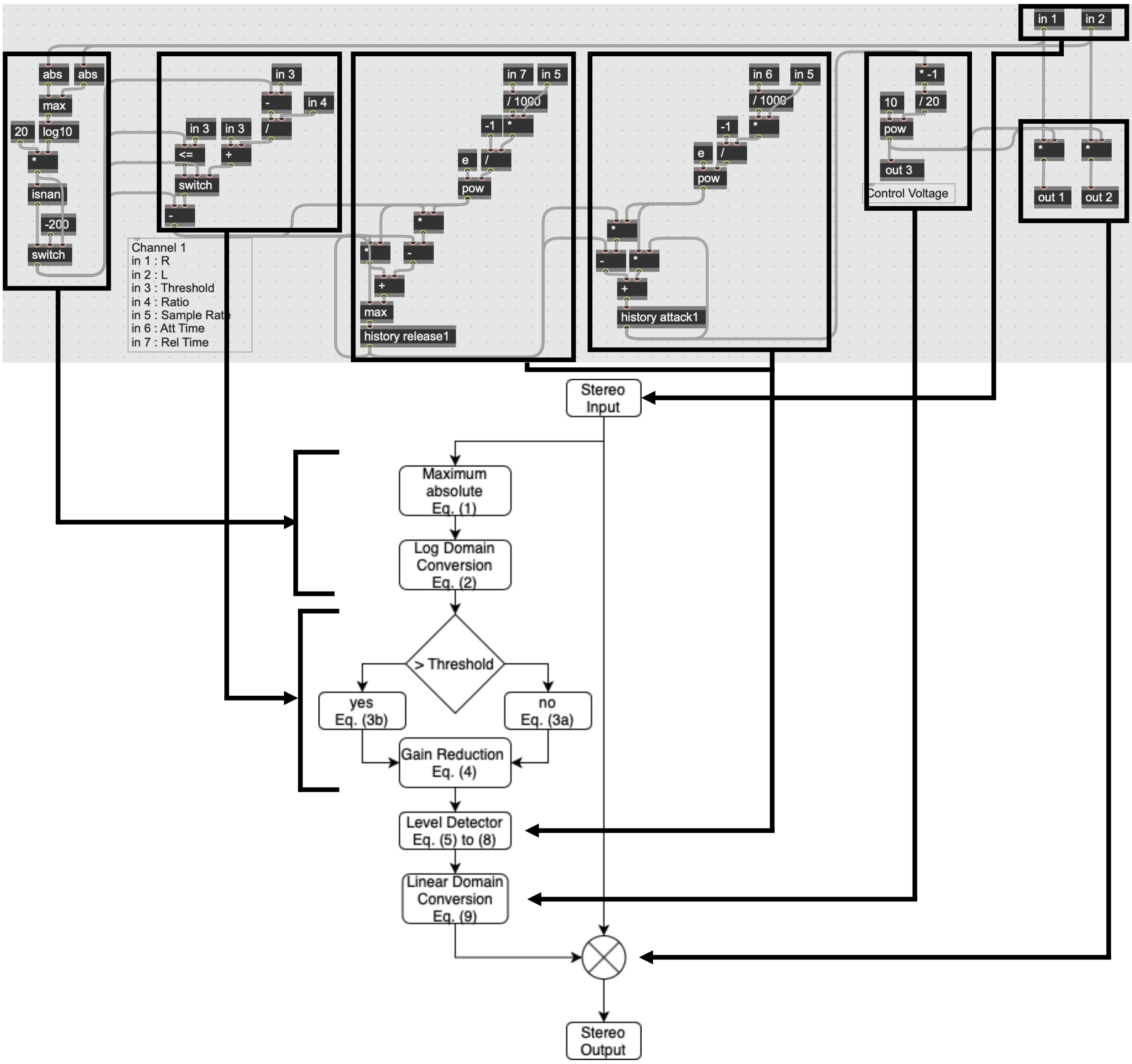


Fig. 3.10 : *Implementation of the auto-adaptive compression architecture in Max 8.*

The second stage is the gain computer. As indicated in the image, ‘in 3’ is the threshold value and ‘in 4’ is the ratio. A Boolean switch is used here to compare the input level to the threshold, whereby if it does not exceed the threshold then the value does not change. However, if the value exceeds the threshold, then it undergoes gain calculation (Eq. 3b). What follows next is the level detection. ‘In 6’ and ‘in 7’ are attack time and release time respectively and ‘in 5’ is the sample rate. The sample rate used for this process is obtained automatically from the audio files loaded to the system using the ‘sfinfo~’ object (not included in this image) that reads and exports the metadata of a sound file. The time parameters are in milliseconds.

Notice that the level detection is divided into two parts and implements two feedback routes one for release and one for attack. This conforms to the architecture of decoupled peak detection (Fig. 2.16). The feedback processing in this section is per sample as opposed to the minimum of 32 sample buffer blocks commonly found in DAWs. The output from this process is then sent to a linear conversion stage and used as the control signal for the compressor. Finally, the control signal is multiplied with the input signal, this is where the audio signal loaded onto the track is either attenuated or unchanged. There are three outputs for this compressor. ‘Out 1’ and ‘out 2’ signify the audio out post-compression, and ‘out 3’ sends out the value of the control signal for monitoring purposes.

3.3.3 Bilateral Compression

Once the auto-adaptive compression has been successfully implemented in Max 8 (Fig. 3.11), creating a bilateral compression system is quite straightforward. The same algorithm in the previous section was duplicated in the same object, modified by adding additional inputs in the gen~ object for Track 2.

Additional inputs for the attack and release time, threshold and ratio for Track 2 were also created to enable independent parameter settings for each track. This creates a feedback cross-adaptivity which was then implemented inside the gen~ object by routing the output of one track to the input of another. Therefore, any changes applied to Track 2 will affect the control signal for Track 1 and vice versa.

Having the algorithms for both tracks housed in one ‘gen~’ object allows for simultaneous data generation and calculation for the two architectures. Using different ‘gen~’ objects for the tracks would introduce unnecessary latency to the process. This is because the output from one ‘gen~’ object will have to be sent to the input of another ‘gen~’ object and vice versa and will cause delay because per-sample transfer between ‘gen~’ object is not possible. The value of the control signal for both tracks can be monitored using the outputs provided. This enables the behaviour of the bilateral compression system to be captured and analysed as part of subsequent testing and experimentation.

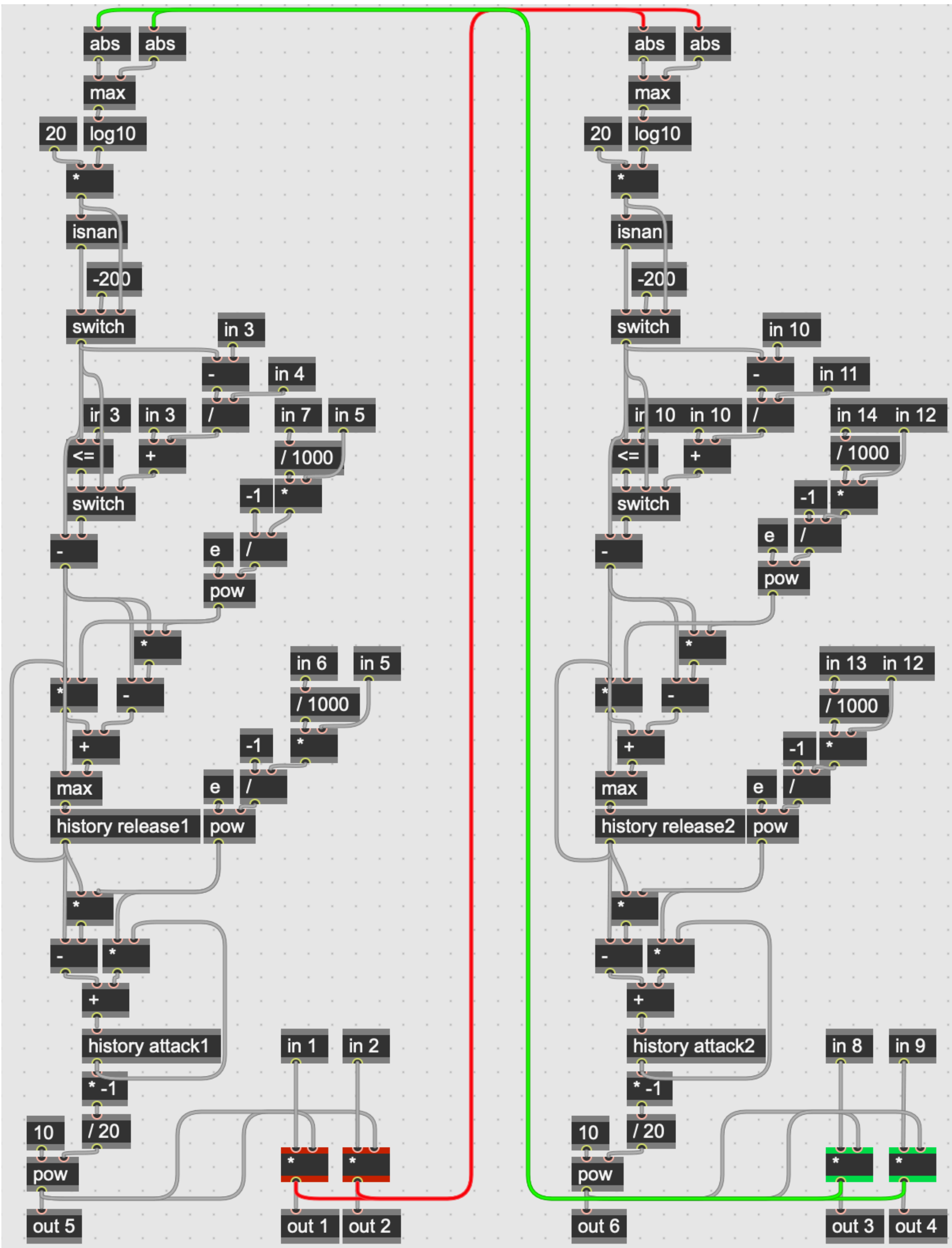


Fig. 3.11 : *The processing architecture for a bilateral compression system in Max 8.*

3.3.4 Compression Curve

Three experiments were conducted to plot the compression curve of the bilateral compression system. To offset the attack and release time, the tests were carried out by raising the input gain of Track 1 by 0.1dB$_{FS}$ increments and collecting the magnitude reading from both outputs (V_{out1} & V_{out2}).

The parameters of the processor were set to values that elicit three different conditions (Eq. 15). The data collected was then plotted in the time domain. The same test signals used in the previous experiments were also used for these experiments to facilitate a comparison.

3.3.4.1 Condition 1

The first condition is when the output for both tracks are in the range that can trigger unidirectional and bilateral compression. The objective of this experiment is to validate the first deduction. The parameters were set to values that are shown in Table 3.3.

As stated in the deduction, the plot should display three distinct sections; two of which represent the result when only one of the outputs exceeds the threshold (unidirectional compression) and one which represents the result when both outputs exceed the threshold (bilateral compression).

Parameter	Track 1	Track 2
Signal Type	Steady-state (-13dB$_{FS}$)	Ramped (-30dB$_{FS}$ to 0dB$_{FS}$)
Threshold (V_{T1} & V_{T2})	-19dB$_{FS}$	-27dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	4:1

Table 3.3 : Parameters for Bilateral Compression Condition 1.

3.3.4.2 Condition 2

The second condition is when both tracks are lower than the threshold. The parameters are given in Table 3.4. The aim of this experiment is to study how the system reacts to these signals.
The Initial deduction is that the graph of this test should not display any changes to the test signals because the input and the output level is inadequate to trigger compression.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-20dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -15dB$_{FS}$)
Threshold (V_{T1} & V_{T2})	-15dB$_{FS}$	-18dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	4:1

Table 3.4 : Parameters for Bilateral Compression Condition 2.

3.3.4.3 Condition 3

The third condition is when both test signals exceed the threshold from the start of the audio. The parameters are given in Table 3.5. Initial deductions suggest that bilateral compression happens instantly until one signal is reduced to values that are lower than its threshold. At this point the behaviour will then change to unidirectional compression with one track becoming the dominant one.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-30dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -0dB$_{FS}$)
Threshold (V_{T1} & V_{T2})	-40dB$_{FS}$	-40dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	10:1

Table 3.5 : Parameters for Bilateral Compression Condition 3.

3.3.4.4 Results

The results from the experiments can be seen in Fig. 3.12 to Fig. 3.14 (pg. 72-74). Noticeable differences can be seen in the plots. However, they conform to the behaviours anticipated earlier.

In the results of the first experiment, the plot (Fig. 3.12) exhibits two knees (marked with arrows) which correspond to the threshold values. Additionally, there are three sections in the graph. Two of them depict cross-adaptive compressions (left and right segments) and one is bilateral compression (middle). It displays stark differences from the result obtained from the previous tests using the DAWs.

The plot for the second experiment (Fig. 3.13) shows no change in the test signals. This is not a surprise as compression does not occur when the magnitude of the signal is lower than their respective thresholds. However, this behaviour demonstrates that the system is functioning correctly in this regard.

The characteristics of the final plot for this section (Fig. 3.14) are as expected whereby mutual compression occurs right from the start. This behaviour changed as soon as Track 1 gets attenuated lower than the value of Threshold 2 (marked with arrow). From then onwards, the behaviour becomes unidirectional sidechain compression where Track 1 is compressed, and Track 2 becomes the dominant. This is because only one of the two signals exceed threshold.

3.3.5 Discussion – Compression Behaviour

The results from these experiments matched the outcome predicted in the previous chapter (Fig. 3.2c, Fig. 3.2d and Eq. 15). For a dynamic signal in bilateral compression system, there are two knees which correspond to the threshold of Track 1 and Track 2. When the magnitude of one signal is equal to or higher than the threshold of the other, the gain of the non-dominant signal is reduced. When both the tracks exceed the thresholds that was set to them, mutual compression happens, and the amount of gain attenuation is given by the formula in Eq. 15.

The compression curve for the bilateral compression system implemented in DAWs, as shown in the previous section, does not resemble the ones presented here. This proves that this architecture can be implemented, but not in DAWs due to either limitations of sample block processing or the lack of cross-modulation compatibility. Now that the behaviour of the bilateral system has been established and validated, the next section examines its behaviour in real-time application. This demonstrates how the system reacts to the time constants.

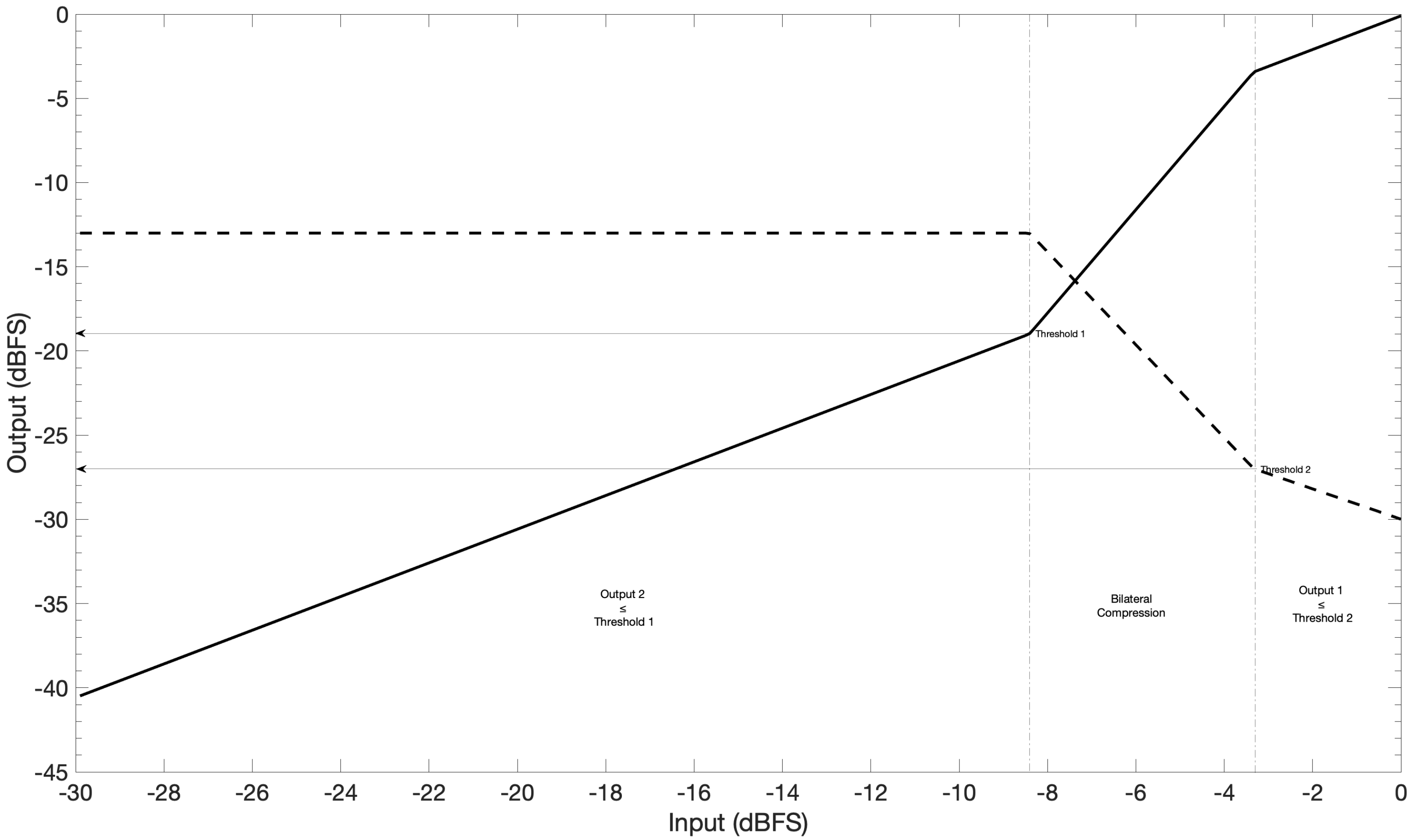


Fig. 3.12 : *Magnitude over time plot of the output from a two-track bilateral compression system. Threshold 1 is -19dB$_{FS}$, Threshold 2 is -27dB$_{FS}$, ratio 1 set to 10:1, and ratio 2 is 4:1.*

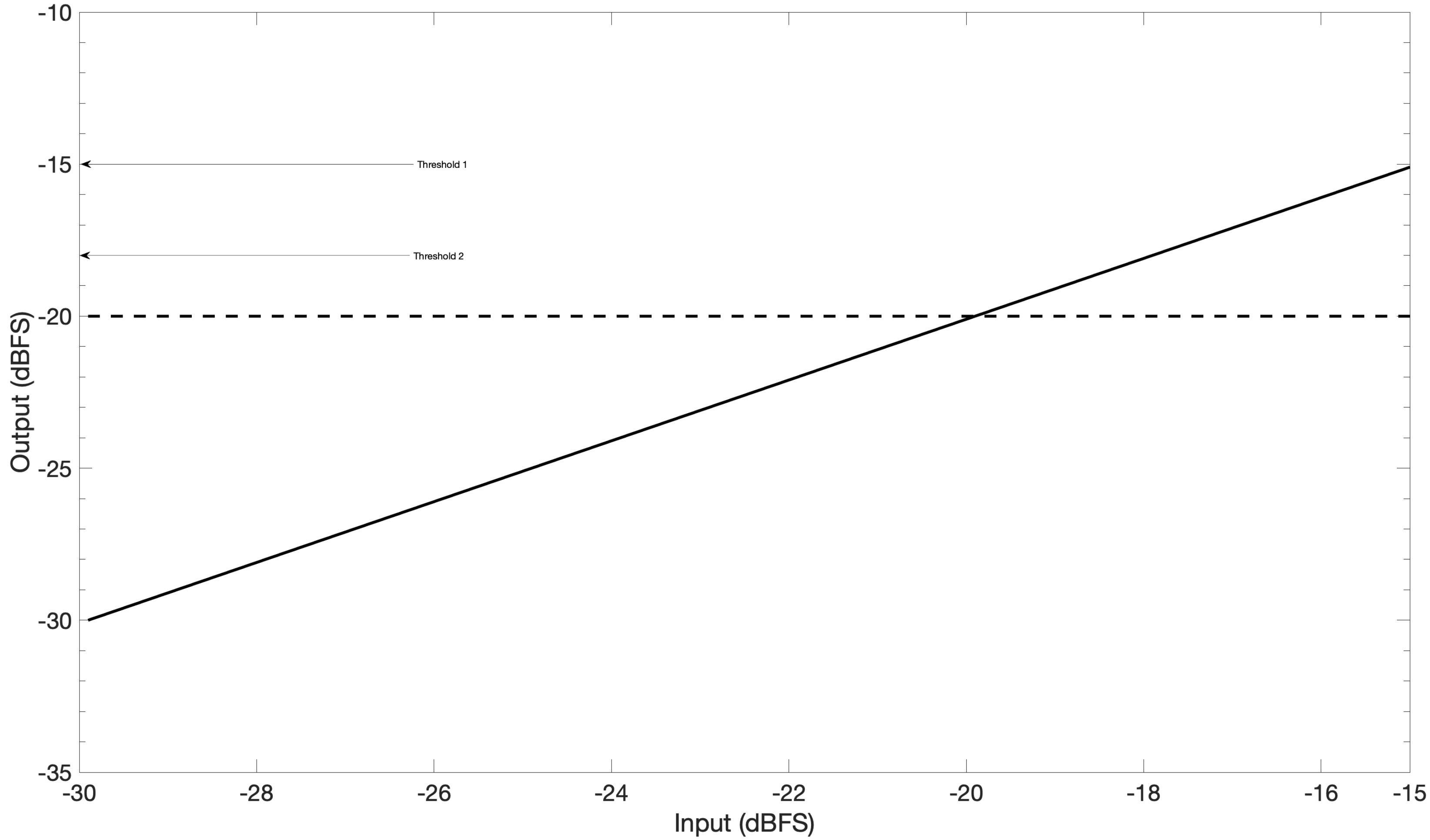


Fig. 3.13 : *Magnitude over time plot of the output from a two-track bilateral compression system. Threshold 1 is -15dB$_{FS}$, Threshold 2 is -18dB$_{FS}$, ratio 1 set to 10:1, and ratio 2 is 4:1.*

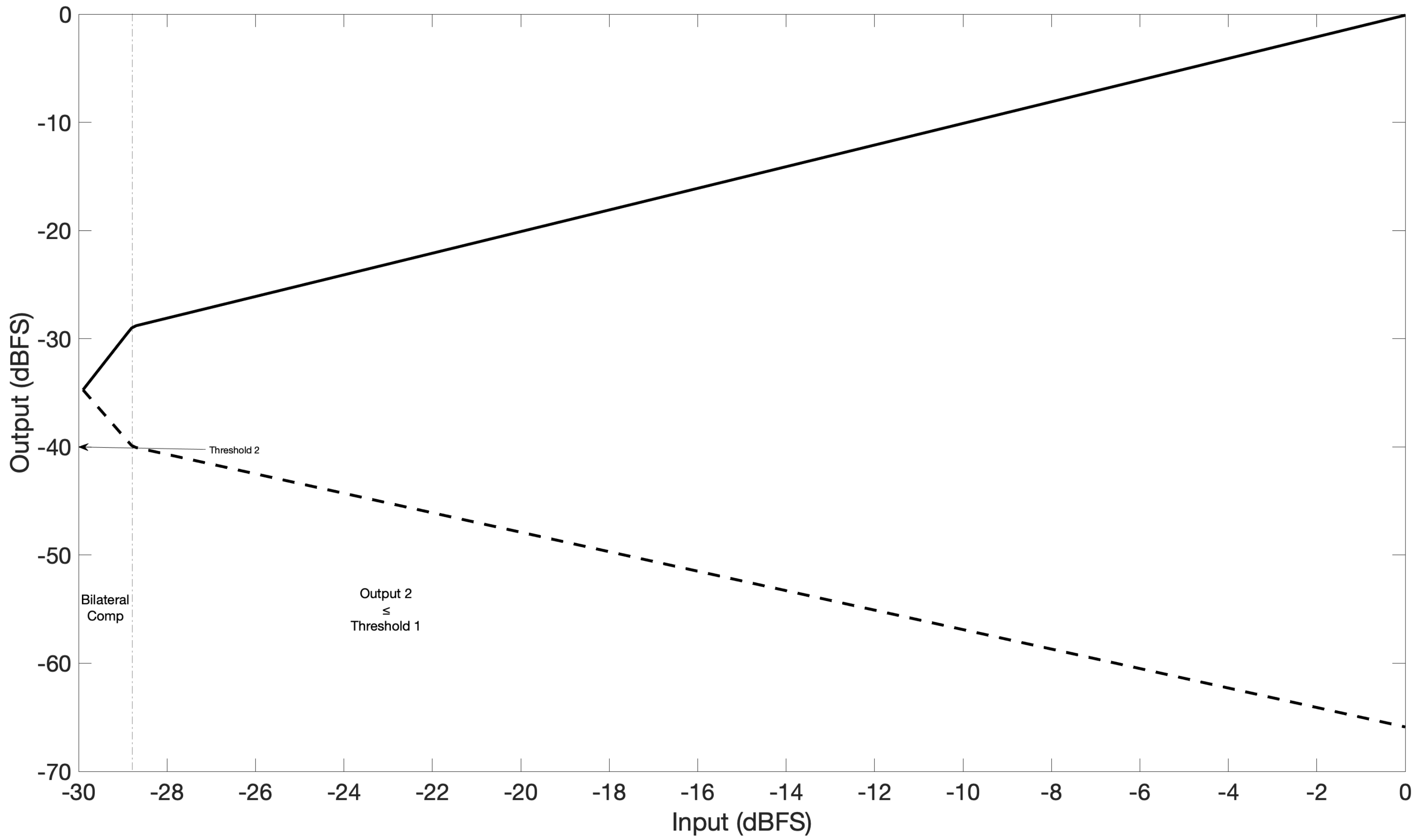


Fig. 3.14 : *Magnitude over time plot of the output from a two-track bilateral compression system, 'Track 1' (dotted) is static -30dB$_{FS}$. 'Track 2' is ramped from -30dB$_{FS}$ to 0dB$_{FS}$, Threshold 1 and 2 is -40dB$_{FS}$,* ratio 1 and 2 is 10:1.

3.3.6 System Behaviour in Real-Time

Another set of experiments was conducted to observe the behaviour of the compressor in real-time, with the attack and release taken into account. The attack and release time chosen for these experiments are suitable for musical material (Wagenaars, Houtsma and Lieshout 1986) with other parameters such as threshold and ratio.

The same test signals from the previous experiment were used for this experiment, using 1kHz sine wave fixed at -13dB_{FS} level for Track 1 and ramped from -30dB_{FS} to 0dB_{FS} in Track 2. Both the signals have a duration of 184 seconds and a sample rate of 44.1kHz. The signals were processed using the bilateral system and the output from both tracks recorded simultaneously in separate audio tracks.

A time domain plot of the magnitude of the output from the tracks was then plotted. The tests aim to investigate how the system responded to the three different conditions. The results from this experiment can be used to compare with the compression curve produced in the previous section.

3.3.6.1 Condition 1

The first test was conducted using different parameter values provided in Table 3.6. The purpose of this experiment was to validate the compression curve produced by real-time application of the system. The parameters used in this experiment are similar to the previous experiment to facilitate comparison.

Based on the compression curve provided in the previous section and the value of the parameters, a few assumptions can be made. The onset of Track 2 will be pushed down by Track 1, whilst the dominant signal is unaffected. There should also be two knees that correspond to the value of the thresholds. The three sections of the plot should also be present.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-13dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -0dB$_{FS}$)
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-19dB$_{FS}$	-27dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	4:1

Table 3.6 : Parameters for Real-Time Bilateral Compression Condition 1.

3.3.6.2 Condition 2

In the second condition, the level of both tracks is lower than the threshold. The parameters are given in Table 3.7. The aim of this experiment is to study how the system reacts to this condition. Just as for the compression curve produced in the previous experiment, if the system performs as it should then both signals should be unaltered.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-20dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -15dB$_{FS}$)
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-15dB$_{FS}$	-18dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	4:1

Table 3.7 : Parameters for Real-Time Bilateral Compression Condition 2.

3.3.6.3 Condition 3

As with the previous section, in the third condition, both test signals exceed the threshold right from the beginning. The parameters are given in Table 3.8. The aim of this experiment is to study how the system reacts to these signals. Based on the compression curve provided earlier, it can be deduced that compression will reduce the level of both tracks within a fraction of a second. As in the graph provided, there will be a point where the system changes from bilateral compression to unidirectional compression.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-30dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -0dB$_{FS}$)
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-40dB$_{FS}$	-40dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	10:1

Table 3.8 : Parameters for Bilateral Compression Condition 3.

3.3.6.4 Result

The plot produced from the first experiment, shown in Fig. 3.15, is identical to the compression curve (Fig. 3.12) indicated by the two knees marked by the arrows. The only difference is that there is an attack slope on the start of the track. That section is enlarged and shown in Fig. 3.16. The plot shows a downwards slope for roughly 5ms before a steep downwards curve. This will be addressed in the discussion. After the 5ms delay, the signal took 10ms to reach roughly 63% of its final value (or $1 - e^{-1}$). This is in accordance with the output value predicted by research cited earlier (Giannoulis, Massberg and Reiss 2012).

For the second experiment, the result (shown in Fig. 3.17) is similar to the plot provided in Fig. 3.13. Compression did not occur because both amplitudes from the output were not strong enough to trigger compression. The result for the third experiment can be seen in Fig. 3.18. It also reflects the graph produced in the previous section (Fig. 3.14).

The only difference is the dip when compression begins during the first few milliseconds of the experiment as can be seen in Fig. 3.18. Significant gain reduction occurred when both the signals were compressing one another excessively during the onset. It took ≈100ms for the signals to reach the intended initial compression value of -34.73dB$_{FS}$. The cause for this phenomenon is still unclear and is under investigation.

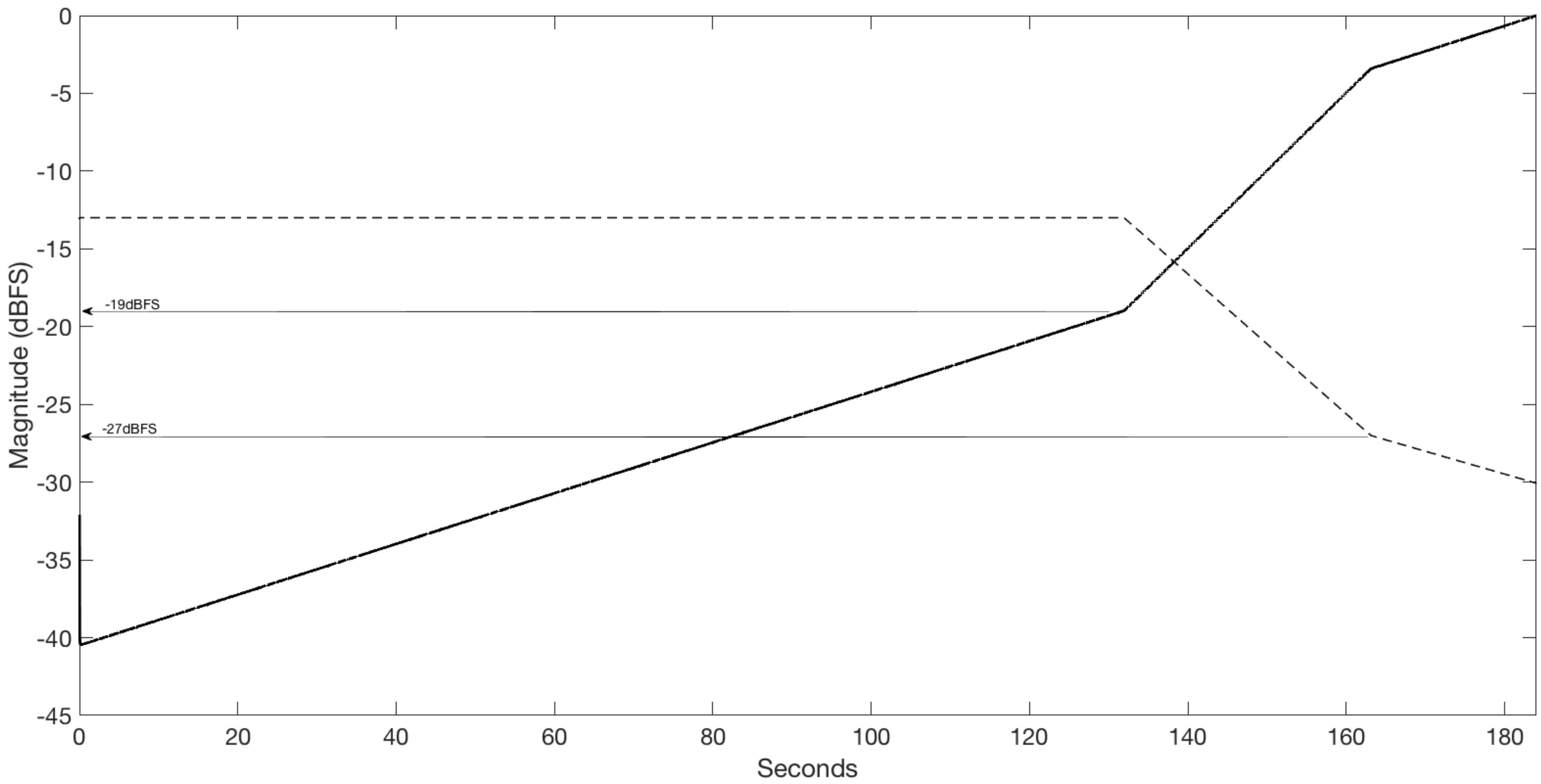


Fig. 3.15 : *Magnitude over time plot of the output from a two-track bilateral compression system. Threshold 1 is -19dB$_{FS}$, Threshold 2 is -27dB$_{FS}$, ratio 1 set to 10:1, and ratio 2 is 4:1.*

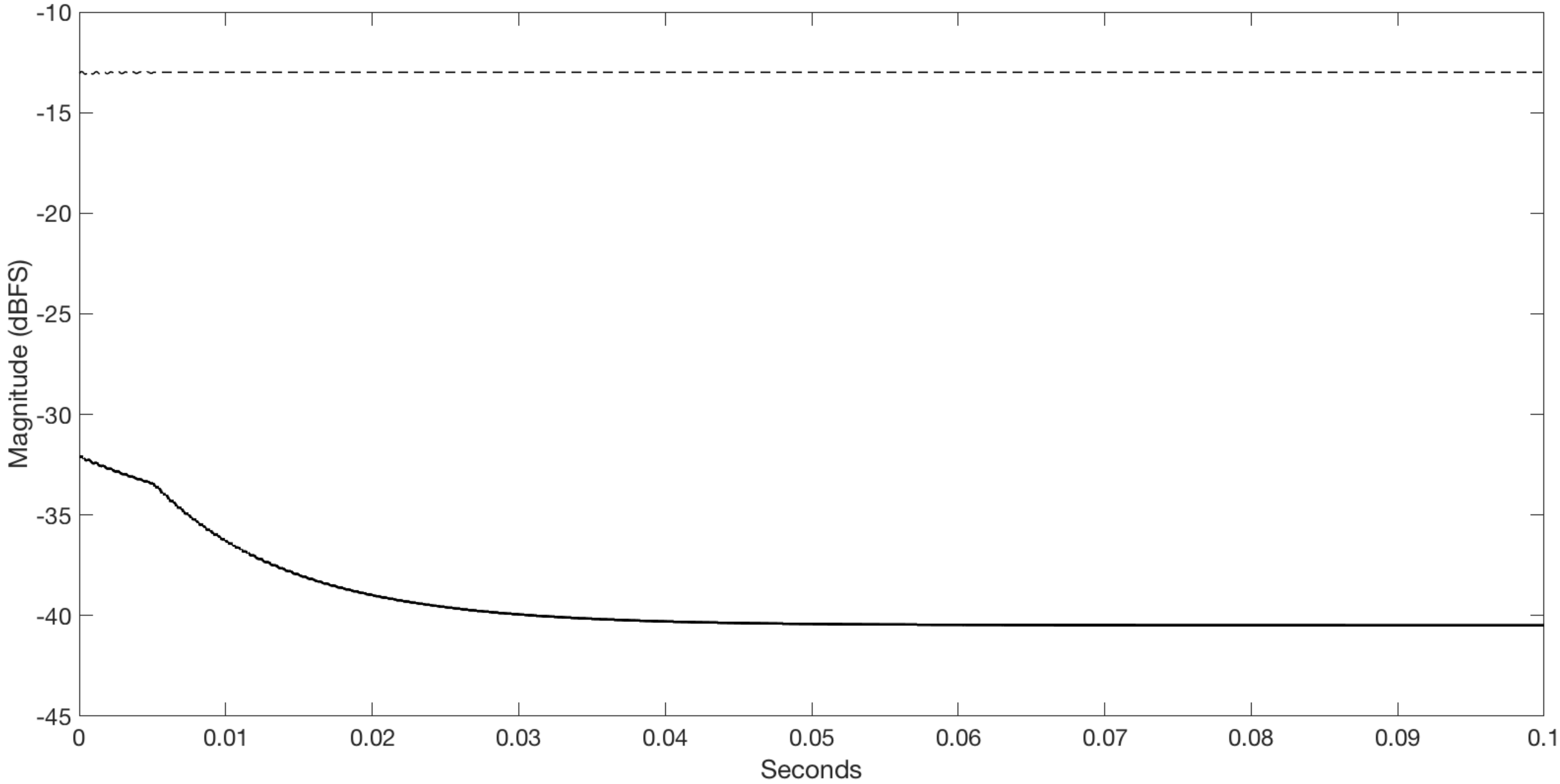


Fig. 3.16 : *Magnitude over time plot of the output from a two-track bilateral compression system, enlarged to 100ms. No compression occurred during the first 5ms.*

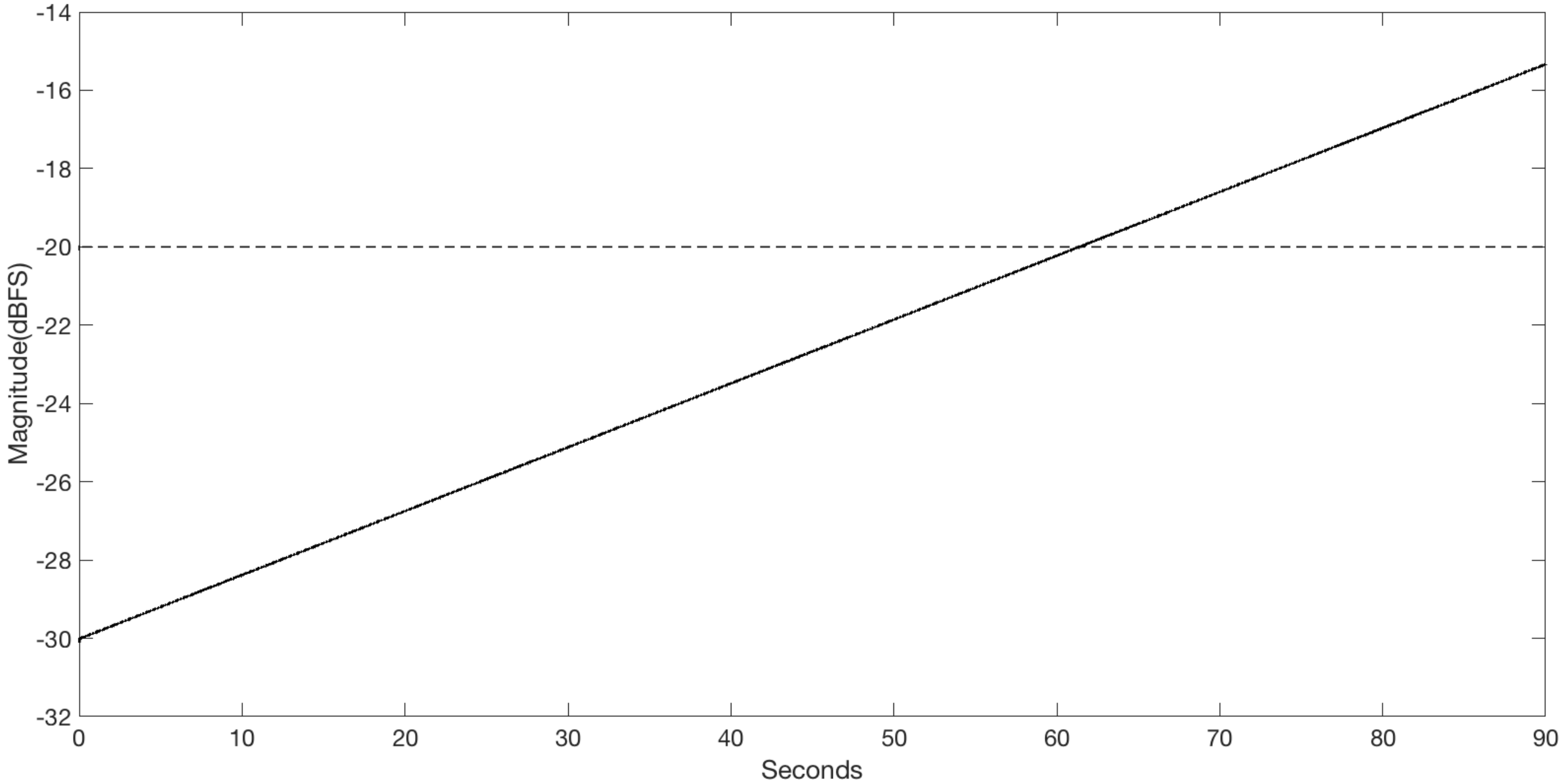


Fig. 3.17 : *Magnitude over time plot of the output from a two-track bilateral compression system. Threshold 1 is -15dB$_{FS}$, Threshold 2 is -18dB$_{FS}$, ratio 1 set to 10:1, and ratio 2 is 4:1.*

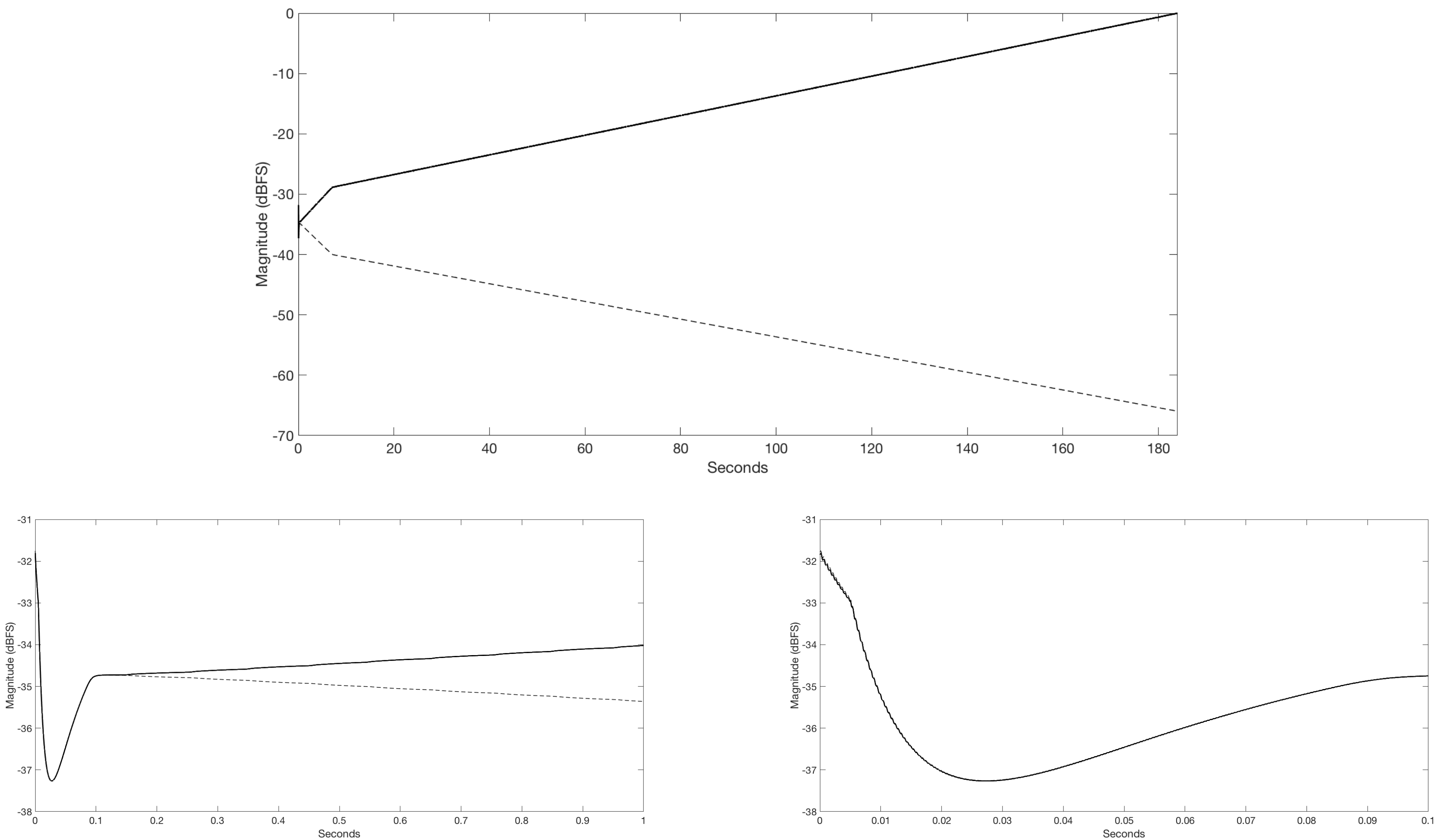


Fig. 3.18 : *Magnitude over time plot of the output from a bilateral compression system set to the same parameters. Track 1 (dotted) is constant -30dB$_{FS}$, Track 2 is ramped from -30dB$_{FS}$ to 0dB$_{FS}$. Output for Track 1 and Track 2; enlarged to 1 second (bottom left), to 0.1 sec (bottom right).*

3.3.7 Discussion – Compression Curve in Real-Time

The result from the real-time application of bilateral compression using test two different test signals produced plots that conform to the predicted compression curve provided in the earlier section. The only difference is in the attack and release of the compression, which were not included in the predictions.

In the plots, there is a constant 0.005 seconds delay before compression occurs. It is likely caused by some kind of start-up delay in the processing. However, this has not yet been confirmed. The phenomenon is still under investigation which subsequently does not affect the signal processing.

Now that the system had been proven to function as intended, more tests were conducted to study how the system processed signals using different parameters, for example, different ratio, and different attack and release time.

3.3.7.1 Condition 4

For this experiment, the objective was to observe the effect of having one track set to a lower ratio compared to the other. The same method where the level for both tracks is higher than the threshold was also used in this test to trigger compression instantaneously.

The goal for this experiment was to examine whether different ratio and different test signals will affect the behaviour of the system. The values of the compressor parameters are given in Table 3.9. The result for when the ramped signal is set to a lower ratio is shown in Fig. 3.19 and for the steady-state signal it is shown in Fig. 3.20.

At the onset of both test signals, the magnitudes are roughly equal. Both of them then undergo gain reduction. The track set to the higher ratio being attenuated more. The tracks then recover from the extreme compression at which point the bilateral compression comes into effect.

Parameter	Track 1	Track 2
Signal Type	Steady-state (-13dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -15dB$_{FS}$)
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms
Threshold (V_{T1} & V_{T2})	-40dB$_{FS}$	-40dB$_{FS}$
Ratio (ρ_1 & ρ_2) - Test 1	5:1	10:1
Ratio (ρ_1 & ρ_2) - Test 2	10:1	5:1

Table 3.9 : Parameters for Real-Time Bilateral Compression Condition 4.

From then onwards, bilateral compression stays in effect until one of the signals is attenuated to values lower than its threshold. Finally, the system behaves like a unidirectional compressor with the signal having the higher amplitude becoming the dominant one.

This is not an unusual characteristic because it is evident that higher ratio will reduce the gain by more. From the compression curve, the behaviour of the system can also be predicted. Both tracks will affect one another because the magnitudes of both signals exceed their thresholds. This continues until the point where one signal is decreased to a value that is not high enough to trigger compression.

To summarise, the ratio of a track affects the compression of a signal in the same way that it would in sidechain compression but with one difference: the amount of gain reduction is governed by the recursive, rather than unilateral, nature of the algorithm.

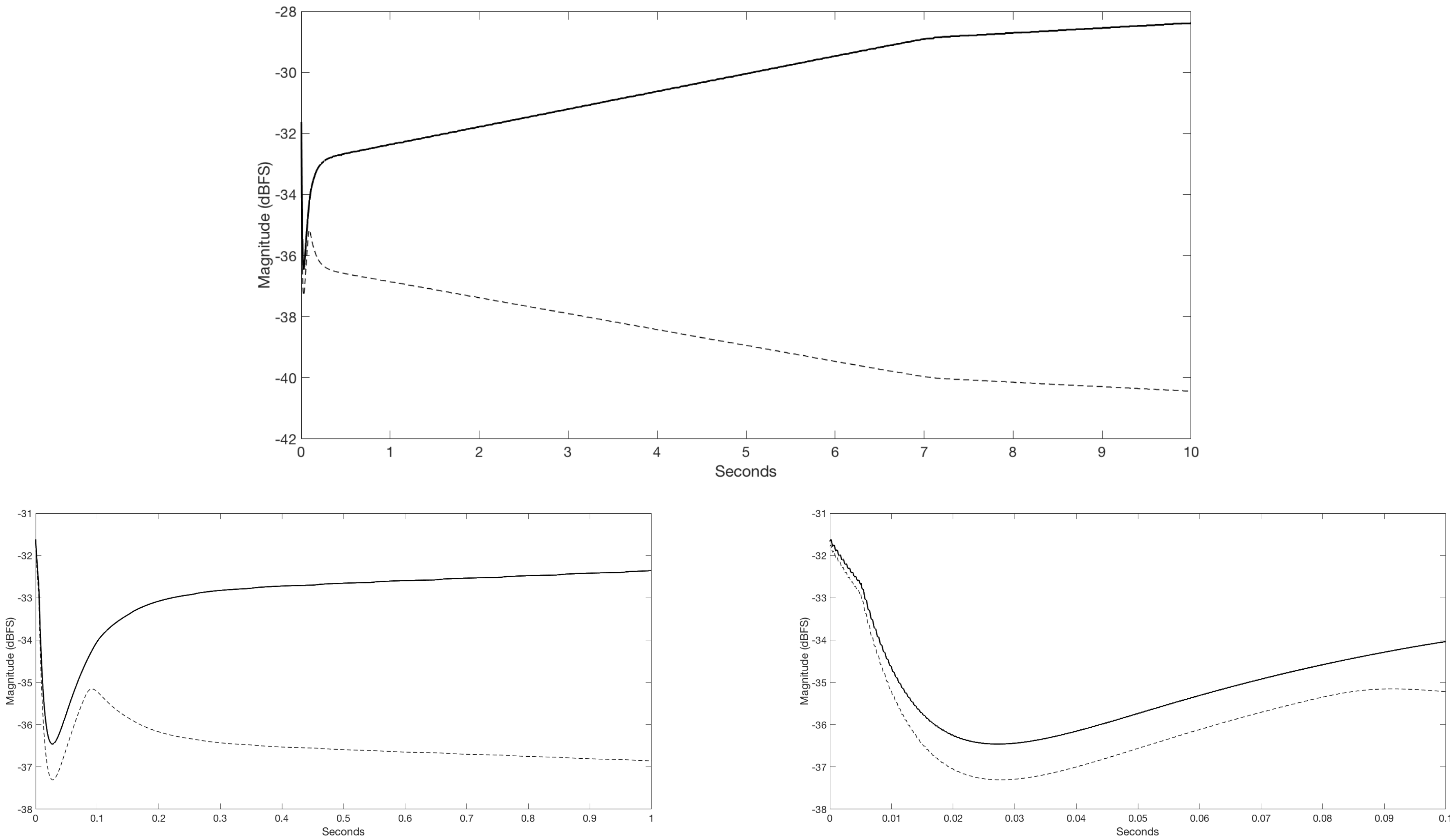


Fig. 3.19 : *Magnitude over time plot of the result from the output of a bilateral compression system using a ramped signal (line) with a compression ratio set lower than the ratio of the static sine (dotted). Top image is the figure enlarged to 10 seconds, bottom left is enlarged to 1 second, bottom right is enlarged to 0.1 second.*

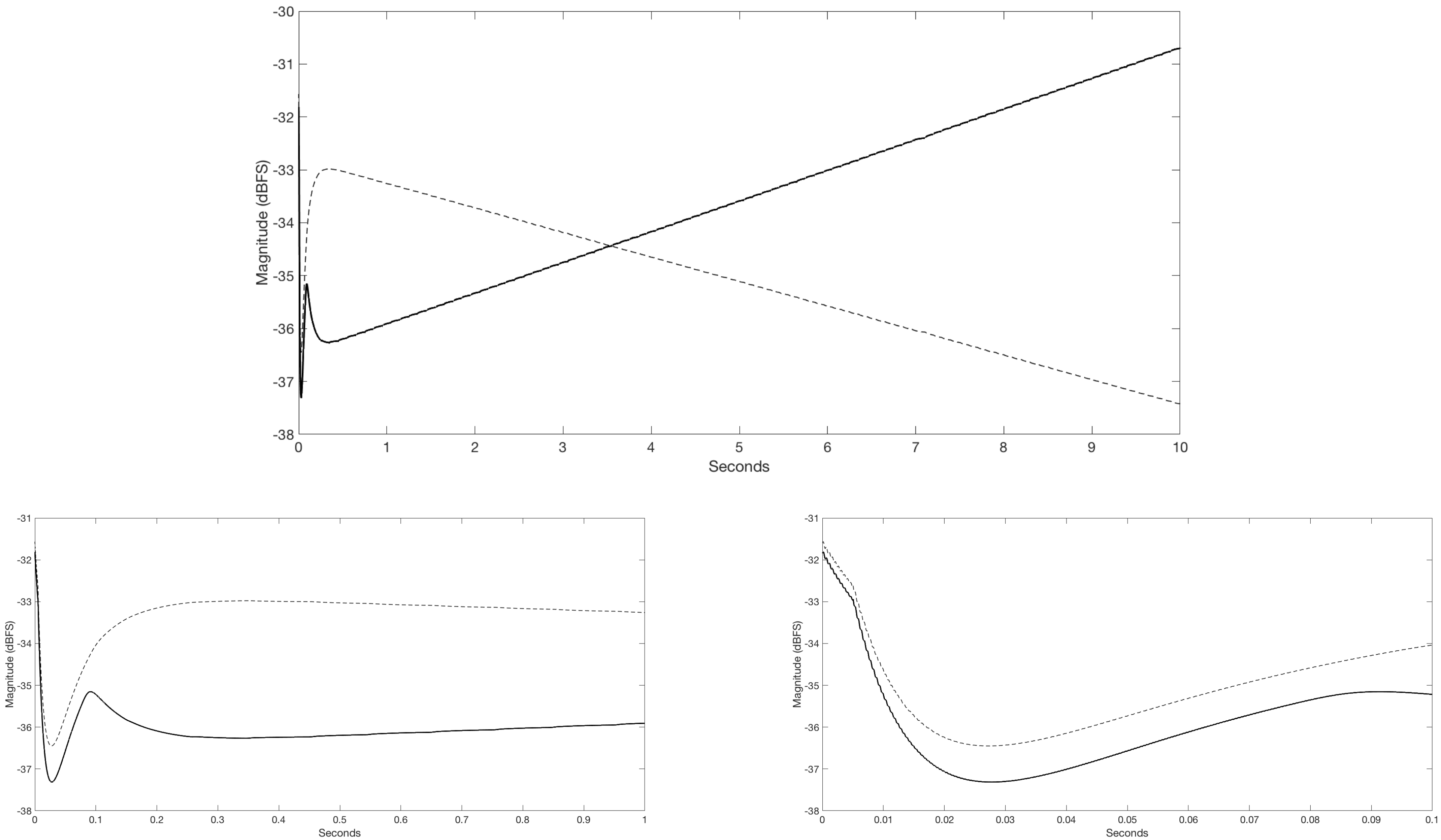


Fig. 3.20 : *Magnitude over time plot of the result from the output of a bilateral compression system using a static sine (dotted) set to a lower compression ratio than ramped signal (line). Top image is the figure enlarged to 10 seconds, bottom left is enlarged to 1 second, bottom right is enlarged to 0.1 second.*

3.3.7.2 Condition 5

The attack and release time for this system might be instrumental in determining a signal's behaviour, the reason being that a longer attack time may lead to a signal having a longer compression onset. This means that if a dominant signal is compressed by another element, it may take a longer duration for it to be fully compressed, leading to a higher chance that it will stay dominant.

A longer release time, however, may cause a signal to stay subjugated for a longer duration. For example, if a signal is in a subjugated state, a longer release time will give it an extended rise time post-compression. This means there is a chance that another signal will compress it in its given state of reduced magnitude.

To investigate the effect of the time constants on a track, an experiment was conducted using the test signals and parameters provided in Table 3.10. The test was conducted twice, where for one track the attack and release time was set to 10ms and 100ms, and to 50ms and 500ms for the other.

Parameter	**Track 1**	**Track 2**
Signal Type	Steady-state (-13dB$_{FS}$)	Ramped (-30dB$_{FS}$ to -15dB$_{FS}$)
Threshold (V_{T1} & V_{T2})	-40dB$_{FS}$	-40dB$_{FS}$
Ratio (ρ_1 & ρ_2)	10:1	10:1
Attack & Release - Test 1	10ms & 100ms	50ms & 500ms
Attack & Release - Test 2	50ms & 500ms	10ms &100ms

Table 3.10 : Parameters for Real-Time Bilateral Compression Condition 5.

The result is shown in Fig. 3.21 and Fig. 3.22. As is generally the case, the track with the slower attack and release time produced a more gradual contour compared to the steep downwards curve produced by the track with quicker attack and release.

One interesting detail that can be observed in the plot is the difference of ≈1.5dB_{FS} in the onset, with the slower attack higher in magnitude. However, the duration is only less than a tenth of a second. Currently this anomaly is unexplained. The result demonstrated the effect of a longer release time when the ramped signal stayed subjugated for a longer duration and the release constant was set to 500ms. A fast attack and release time also made the signal more dynamic. The signal's magnitude fluctuates within a short duration.

3.3.8 Discussion – System Behaviour in Real-Time

The first section has not only proven that the behaviour of this system adheres to its theoretical construct, but also demonstrated a general compression curve for a bilateral compression system, processing a steady-state and a ramped signal. Furthermore, the compression curve also reveals how ratio and time constants affect the behaviour of the system. To investigate whether this architecture fulfils the second criteria - the ability to perform cross-modulation - another set of experiments was conducted using amplitude modulated (AM) signals (Reiss and Brandtsegg 2018). AM signals were chosen because they provide a visual indication of the effect of cross-modulation in the amplitude and frequency over time plots.

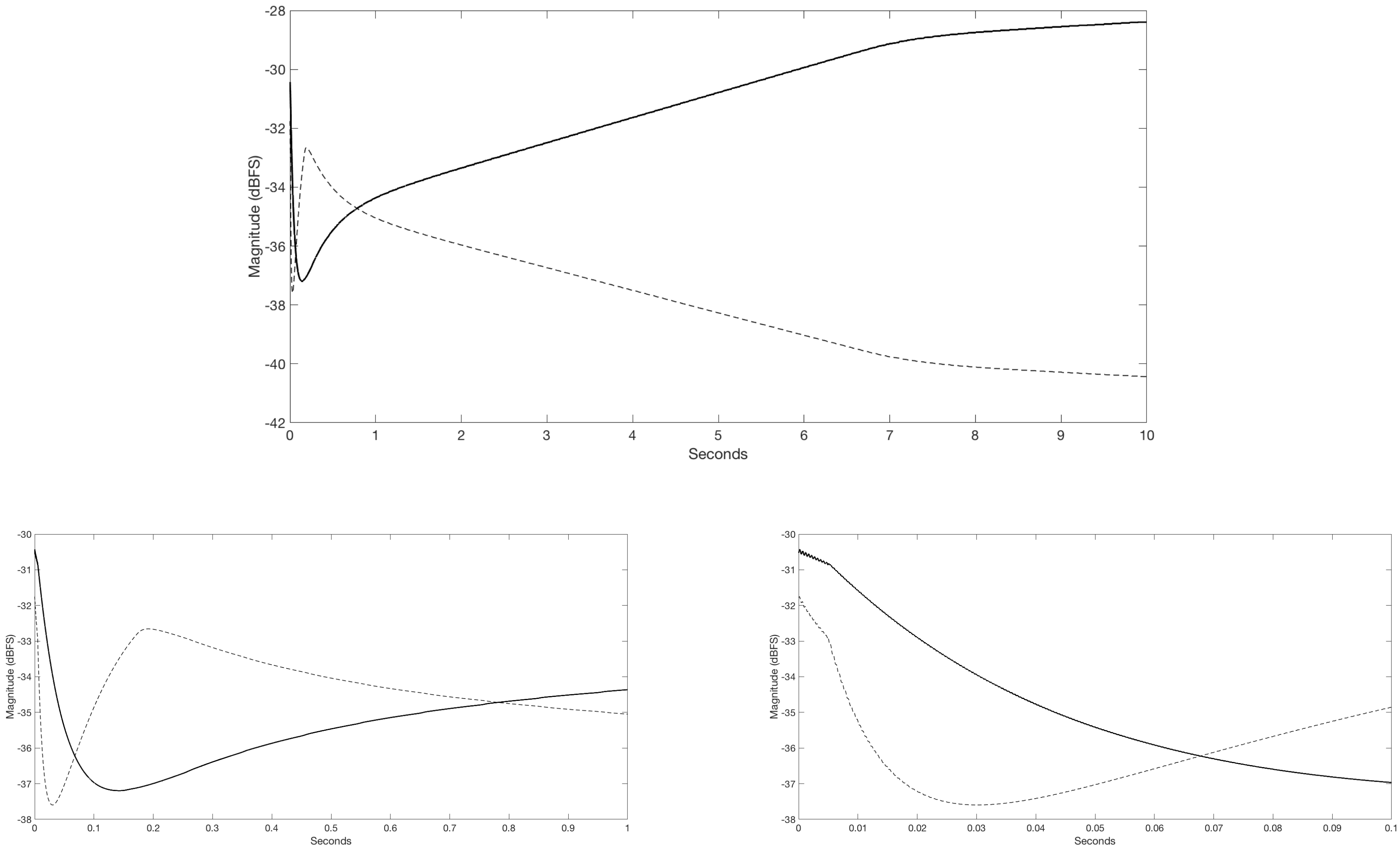


Fig. 3.21 : *Magnitude over time plot of the output from a bilateral compression system with the ramped signal (line) set to a slower attack and release time compared to the static sine (dotted). Top image is the figure enlarged to 10 seconds, bottom left is enlarged to 1 second, bottom right is enlarged to 0.1 second.*

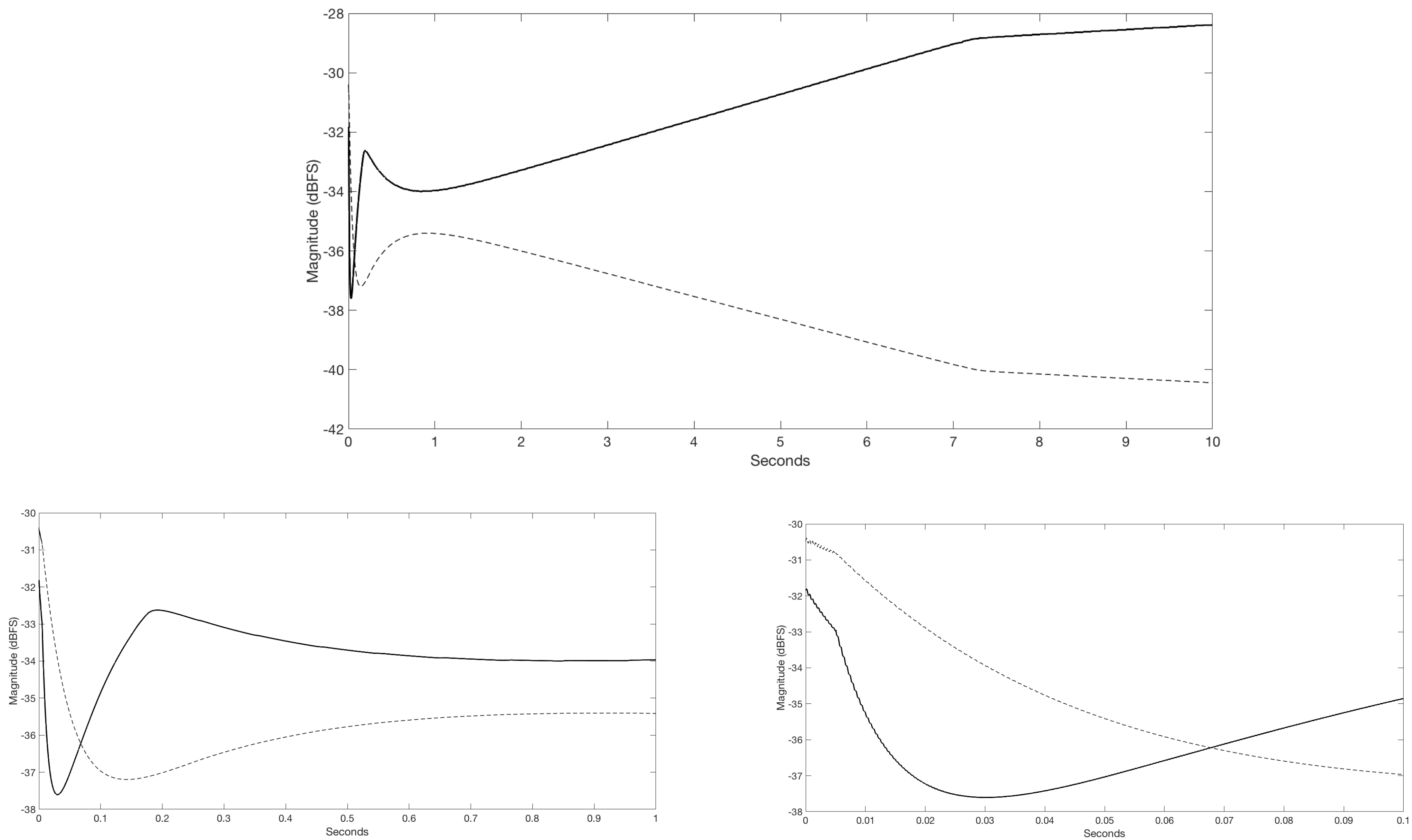
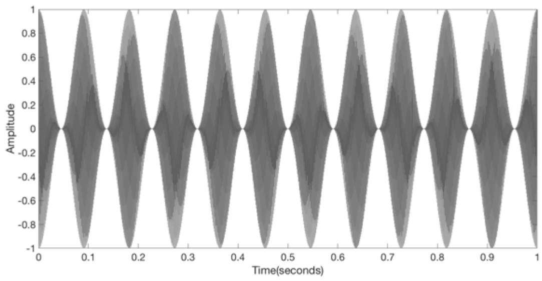


Fig. 3.22 : *Magnitude over time plot of the output from a bilateral compression system with the steady-state signal (dotted) set to a slower attack and release time compared to the ramped signal (line). Top image is the figure enlarged to 10 seconds, bottom left is enlarged to 1 second, bottom right is enlarged to 0.1 second.*

3.3.8.1 Amplitude Modulated Signal Test

This experiment aims to verify the occurrence of cross-modulation in the bilateral compression system created in Max 8. Furthermore, this will enable further investigation into how one signal will affect another signal by means of imprinting its characteristics. Tests were conducted using a pair of sine waves (see Fig. 3.23) modulated by two subsonic (i.e., less than 20Hz) modulators of proximate primes. The test signals were 6 seconds in duration with sample rate of 44.1kHz and were normalised to 0dB$_{FS}$.



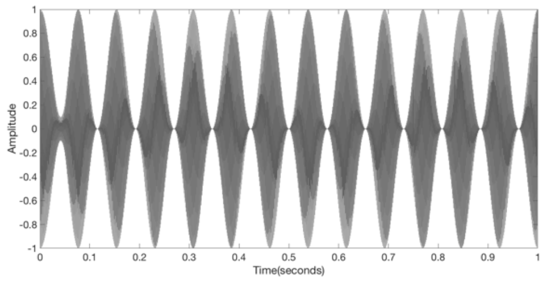


Fig. 3.23 : *Amplitude over time plot of the two test signals; a 997Hz sine wave modulated by 11Hz (left) and 13Hz (right) sine wave.*

Primes were chosen to reduce the possibility of artefacts of the processing occurring at the same position on the spectrum. Subsonic modulated test signals were used so that the effects of the compression could be observed as additional spectral components around the carrier frequencies that the modulators were applied to. The parameters for this experiment are provided in Table 3.11. The compressor in this test applies a low threshold with a high ratio to strongly exercise the compressor. This is because when compression is used in an obvious way, the artefacts are typically heard in the time domain (i.e., noise pumping). The attack and release time were chosen so that compressor action would only affect (i.e., distort) the waveform of the modulator without changing the shape of the carrier wave.

The experiment was initially undertaken with one signal more dominant to investigate how the dominant signal affected the weaker one. This was achieved by lowering the magnitude of one track in the system by 6dB. A larger difference in magnitude might cause too much compression on the non-dominant test signal, eliminating the possibility of observing any interactions between the two. The final experiment was done with the signal having the same level of 0dB$_{FS}$.

Parameter	Track 1	Track 2
Signal Type	997Hz x 11Hz sine wave	997Hz x 13Hz sine wave
Threshold (V_{T1} & V_{T2})	-30dB$_{FS}$	-30dB$_{FS}$
Ratio (ρ_1 & ρ_2)	20:1	20:1
Attack (α_A)	10ms	10ms
Release (α_R)	100ms	100ms

Table 3.11 : Parameters for Amplitude Modulated Signal Test.

3.3.8.2 Result – Time Domain

For the first experiment, Track 1 was made the dominant track in the system. By examining the time domain plots of the output (Fig. 3.24), one can conclude that the dominant signal is also affected by the subjugated signal. This is indicated by the peaks and dips on both plots that coincide with one another. For instance, in between the 0.7 and 0.78 second mark in Track 1 there is a reduction in amplitude. At the same point in Track 2, the gain reduction is not as intense as compared to the peak before and after the 0.7 to 0.78 second mark. Referring to Fig. 3.23, at the point in question, the modulation frequency cycle for Track 1 has just started, while for the other track it is at its peak. Therefore, the magnitude for the 11Hz signal is not strong enough to trigger compression.

For the second test, Track 2 was made the dominant element. Similar behaviours can be seen in the time domain plot (Fig. 3.25) where the amplitude for the dominant signal is attenuated when there is a peak on the non-dominant signal. This shows that the interplay between both signals for this experiment is dependent on the difference of amplitude caused by the modulation frequency cycle, implying cross-modulation. One interesting detail highlighted by the results from this section is that the attenuation of the dominant signal in Fig. 3.25 is not as powerful as in the first experiment. This may be because the dominant signal has a shorter modulation cycle (i.e.,13Hz), enabling it to recover from compression quicker.

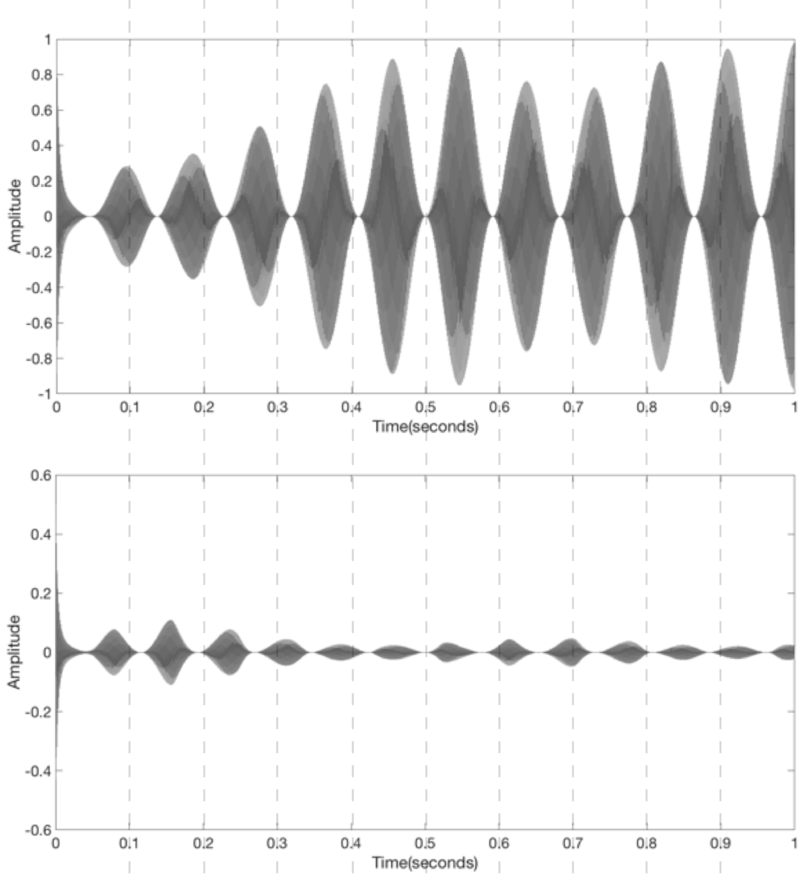


Fig. 3.24 : *Amplitude over time plot of the 11Hz modulated test signal (top) and the 13Hz modulated test signal (bottom) in a bilateral compression system with the 11Hz signal 6dB higher than the 13Hz.*

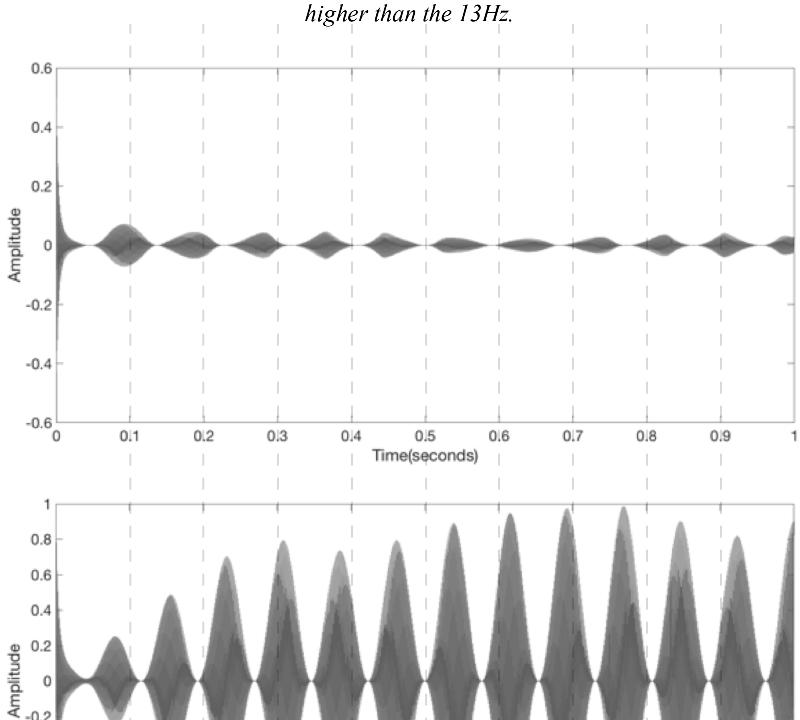


Fig. 3.25 : *Amplitude over time plot of the 11Hz modulated test signal (top) and the 13Hz modulated test signal (bottom) in a bilateral compression system with the 13Hz signal 6dB higher than the11Hz.*

The time domain plot for when both test signals have the same loudness highlights clear coaction (Fig. 3.26). The peaks and attenuation are complementary to one other, which relates to the cycle of the modulation frequency. This further strengthens the assumption of cross-modulation in this system. Thus, to verify this assumption, frequency magnitude plots of the result from these tests were produced.

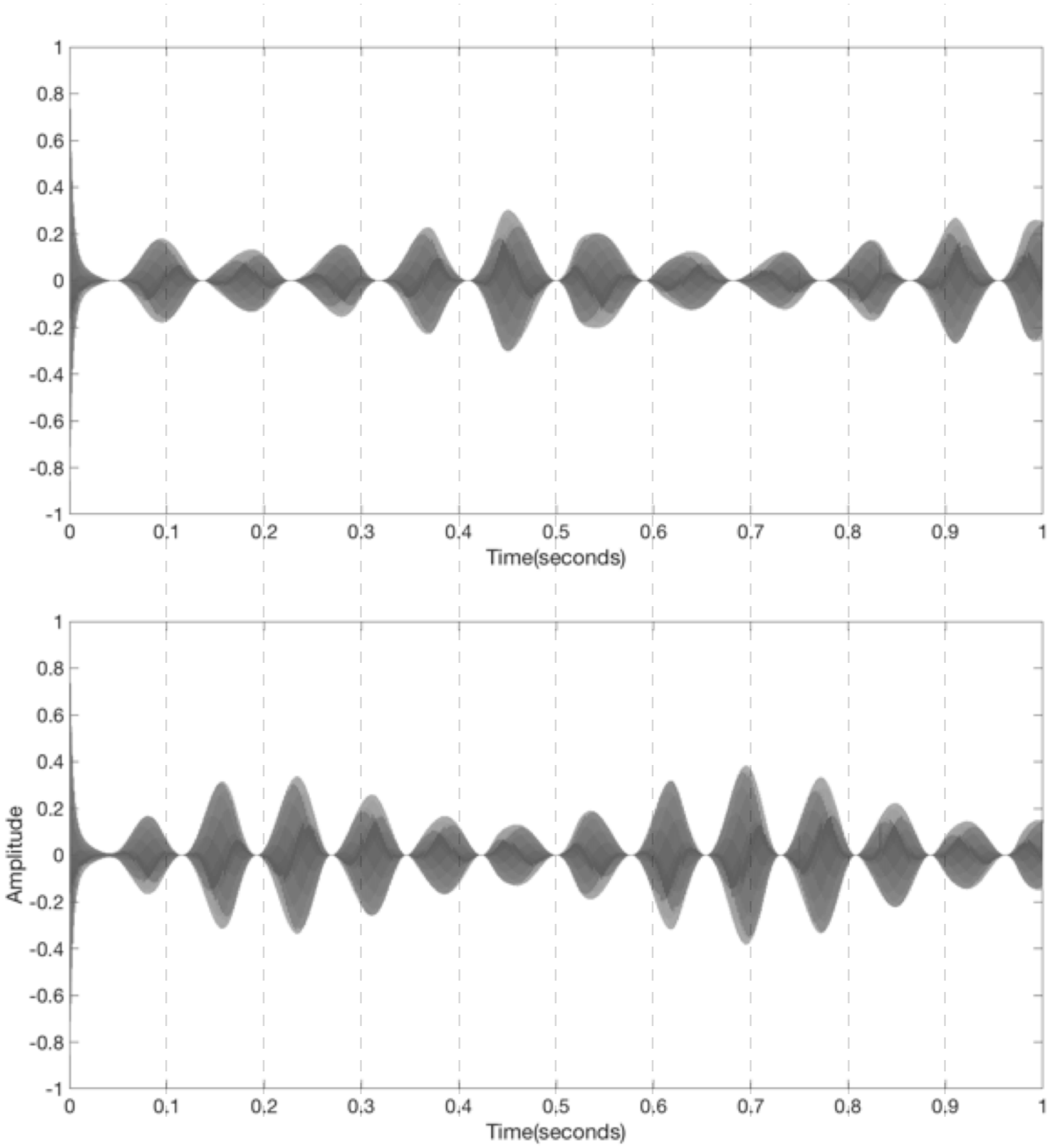


Fig. 3.26 : *Amplitude over time plot of the 11Hz modulated test signal (top) and the 13Hz modulated test signal (bottom) in a bilateral compression system with both signals set to* $0dB_{FS}$.

3.3.8.3 Result – Frequency Domain

Frequency magnitude plots of the results from the previous tests were produced to aid in validating the occurrence of cross-modulation in the architecture. However, to facilitate a better understanding of the plots and the characteristics of cross-modulation, a comparison of the frequency magnitude graphs generated from the adaptive compression (i.e., self-compression), cross-adaptive compression (i.e., sidechain compression) and also bilateral compression were conducted.

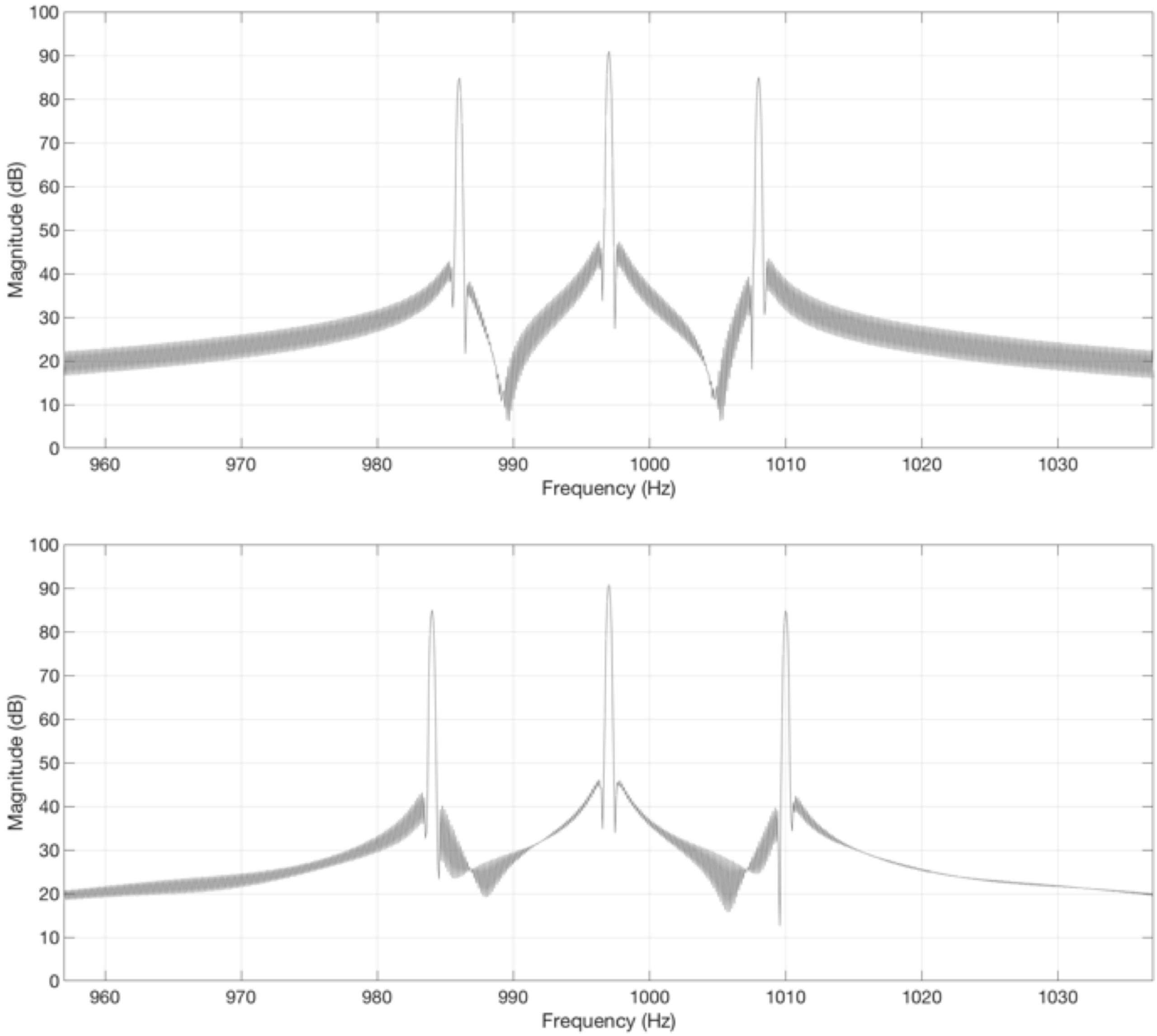
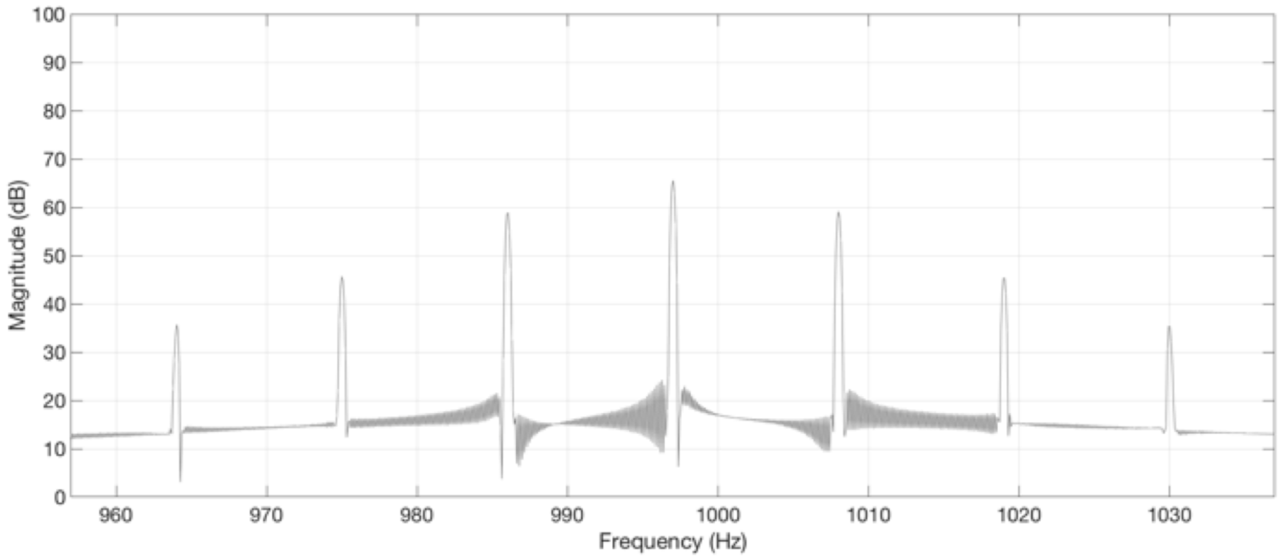


Fig. 3.27 : *Frequency magnitude plot of the two test signals; a 997Hz sine wave modulated by 11Hz sine wave (top) and a 997Hz sine wave modulated by 13Hz sine wave (bottom).*

First, this thesis presents the charts for the two test signals (Fig. 3.27). The sum and difference components from the modulation are apparent in the figure. The three peaks in the plot are the carrier frequency (middle - 997Hz) and the sum and difference element from the modulation which are 11Hz and 13Hz deviation from the carrier frequency.

Next is the result of auto-adaptive compression. For this sample, the parameters of the compressor were set to;

Threshold = -30dB_{FS}, ratio = 20:1, attack time = 10ms and release time = 100ms.



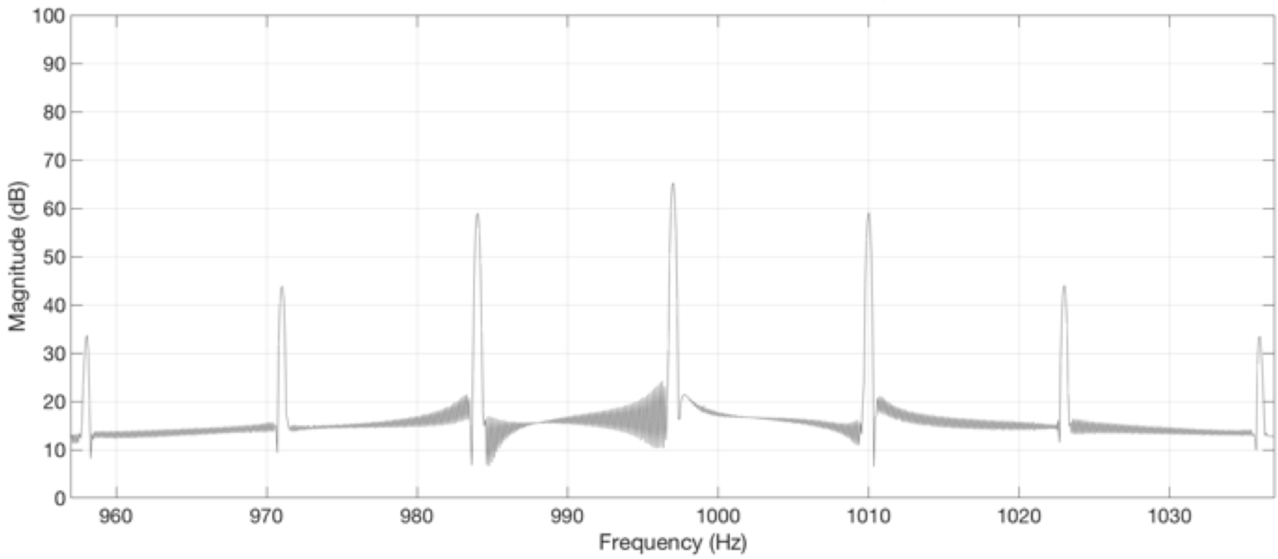


Fig. 3.28 : *Frequency magnitude plot of the result from the auto-adaptive compression test; the 11Hz modulated test signal (top) and the 13Hz modulated test signal (bottom).*

As shown in Fig. 3.28, new spectral content was introduced. More than three peaks can be seen in the figure which correlate to the sum and difference elements of the test signals. The frequency difference between every peak in Track 1 are 11Hz and they are 13Hz for Track 2.

For sidechain compression, the signal is compressed using the values calculated by another track. Thus, not only is the signal compressed, but it also has the character of another signal imprinted on to itself. Therefore, the graph should exhibit both the self-modulation sidebands and sidebands imparted by the dominant track. The dominant signal, however, should be unaltered.

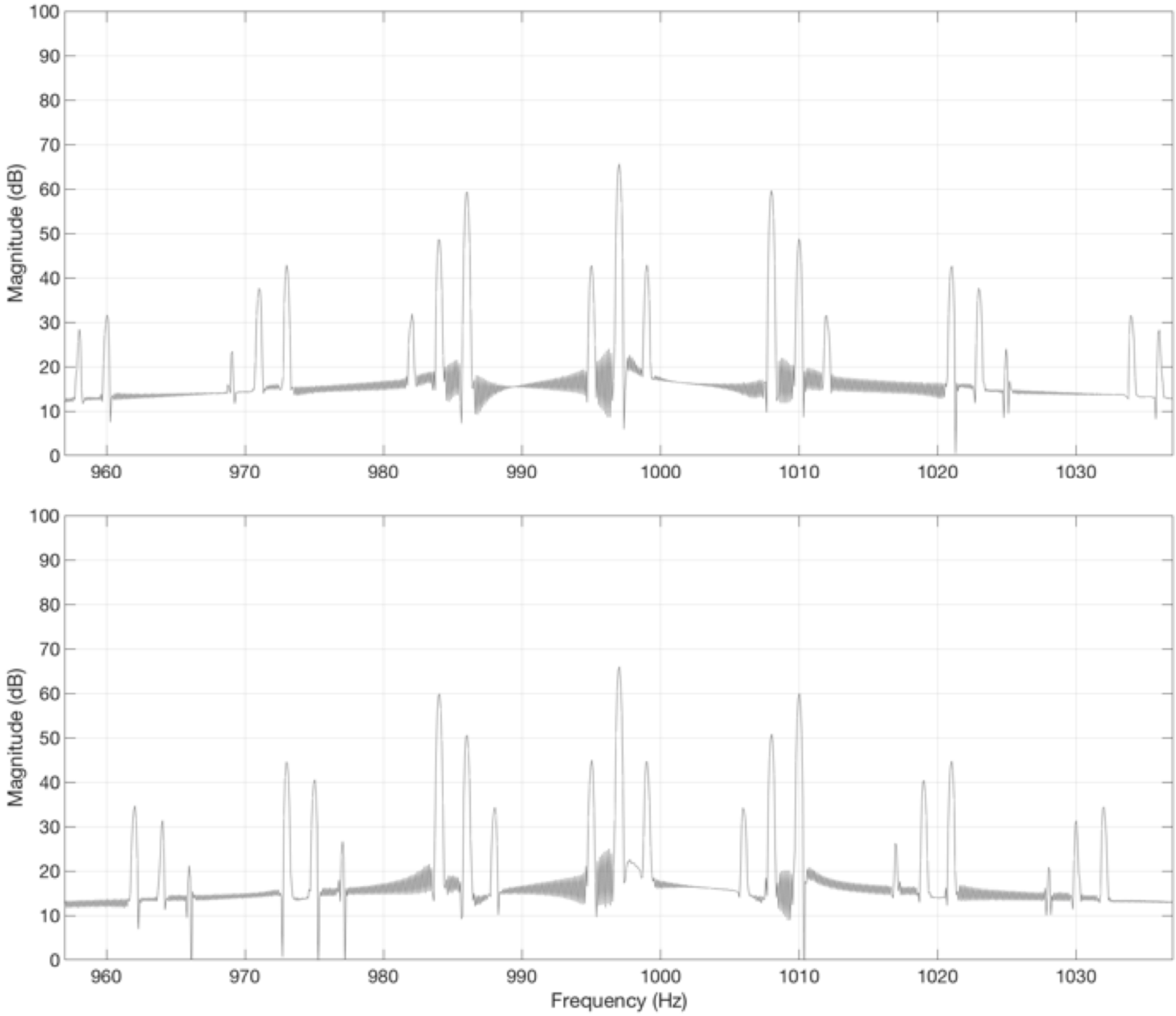


Fig. 3.29 : *Frequency magnitude plot of the result from the cross-adaptive compression test; the 11Hz modulated sine wave attenuated by the 13Hz modulated sine wave (top) and the13Hz modulated sine wave attenuated by the 11Hz modulated sine wave (bottom).*

The frequency magnitude plot of the sidechain architecture (Fig. 3.29) shows more sidebands compared to the previous graphs. Considering that the sidechain architecture is unidirectional, the dominant signal is not affected by the subjugated signal and therefore the plots are the same as Fig. 3.27.

The new spectral content of the subjugated signals corresponds to the sum and difference of themselves and the dominant signal. The additional sidebands are of 2Hz deviance from the original spectral content generated by self-compression (Fig. 3.28). The magnitude of the new sidebands, however, are reduced every iteration.

Based on the results, a deduction can be made from the plot of the bilateral compression. If for a sidechain system, the spectral content consists of the self-compression sidebands and the dominant signal sidebands, then a bilateral compression system, owing to its recursive nature, will contain sidebands of sidebands that get progressively weaker.

Now that the spectral behaviour of self-compression and sidechain compression has been studied, the plot of the results from the bilateral compression test are provided to validate the occurrence of cross-modulation. When one signal in the system is dominant, its spectral identity (i.e., the three original peaks) is more prominent (Fig. 3.31a and Fig. 3.31b). There is also novel spectral content 2Hz apart from one another however the magnitude is low.

For the subjugated track however, the magnitude of the three initial peaks is now reduced. Furthermore, the magnitude is similar to the additional sidebands. For the conditions where both test signals have the same magnitude, new spectral components are prominent (Fig. 3.30). The magnitude of the new elements is not much different from the carrier frequency and the initial sidebands.

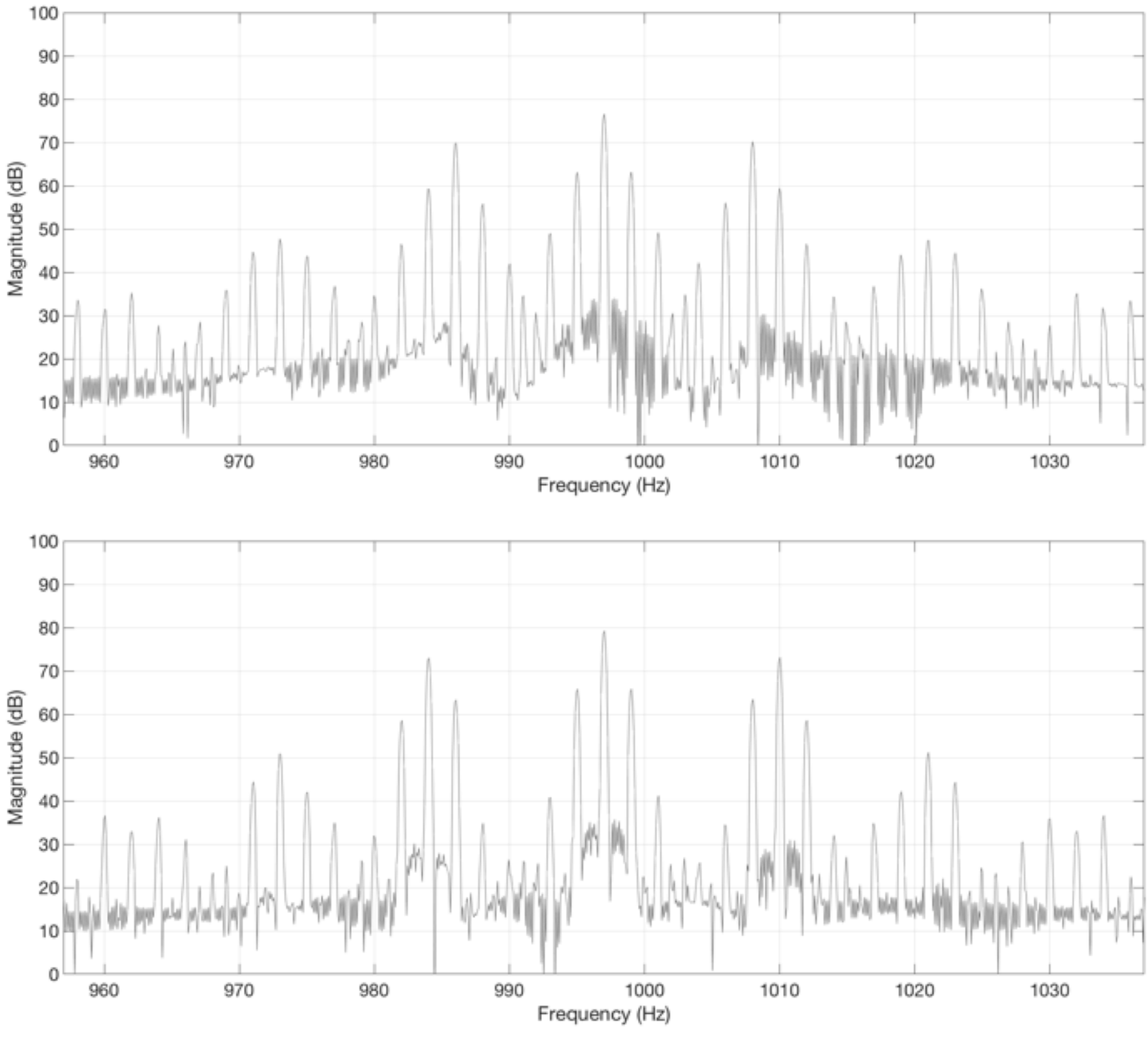


Fig. 3.30 : *Frequency magnitude plot of the result from the bilateral compression test; the 11Hz modulated sine wave (top) and the 13Hz modulated side wave (bottom). Both test signals were set to 0dB$_{FS}$.*

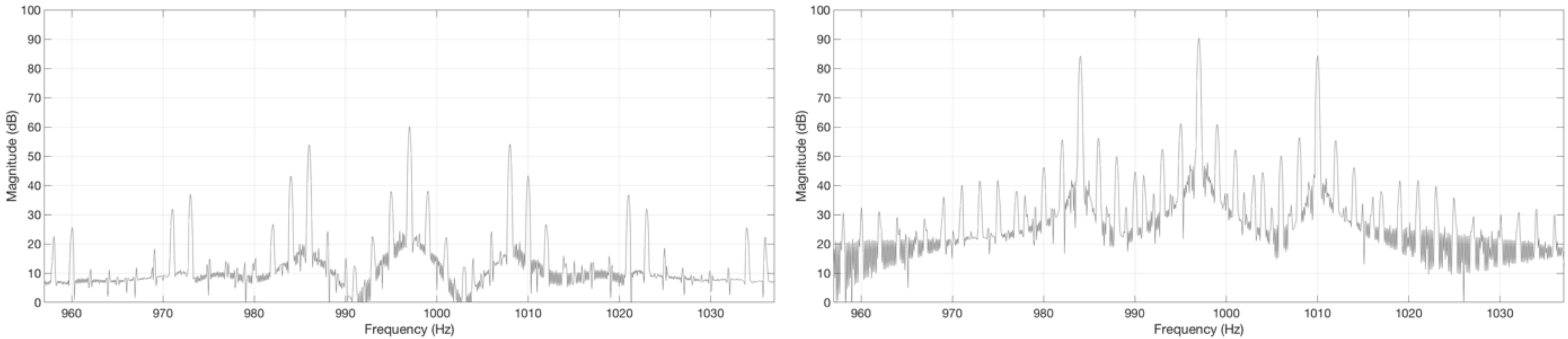


Fig. 3.31a : *Frequency magnitude plot of the result from the bilateral compression test; the 13Hz modulated sine wave dominant (right) subjugating the 11Hz modulated sine wave (left).*

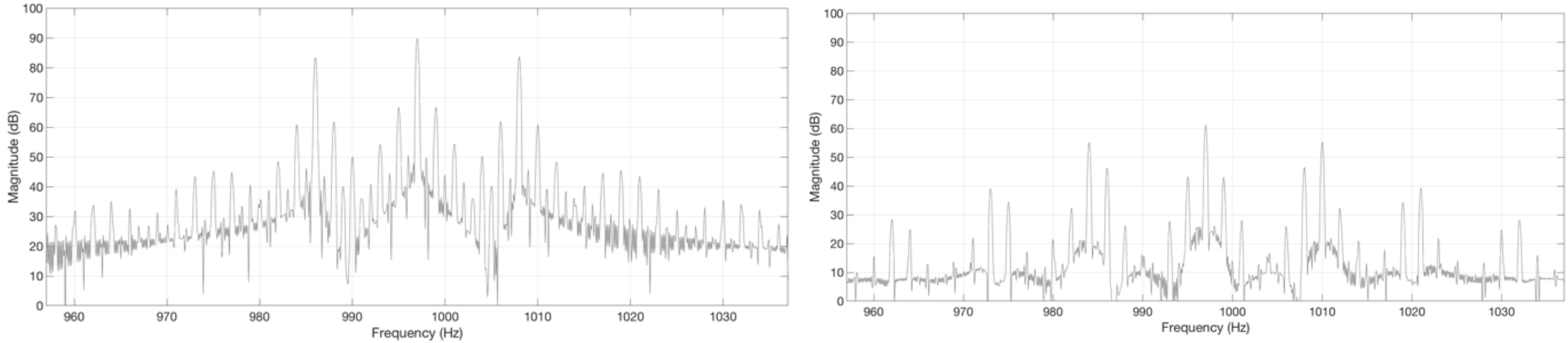


Fig. 3.31b : *Frequency magnitude plot of the result from the bilateral compression test; the 11Hz modulated sine wave dominant (top) subjugating the 13Hz modulated sine wave (bottom).*

3.3.8.4 Condition 6

Now that the occurrence of cross-modulation in this system has been verified, there is one more behaviour left to be explored. So far, steady-state sounds have been used to compress one another, as well as AM signals. However, musical compositions are constructed by a myriad of instruments with varying timbres often playing at the same time. This is addressed via the creation and discussion of musical examples in the next chapter.

Imagine the test signals used are a steady state 1kHz sine wave at $0dB_{FS}$ and a 907Hz sine wave modulated by a 3Hz sine. The AM sine wave is ramped from $-60dB_{FS}$ to $0dB_{FS}$. The threshold for both tracks is set to $-10dB_{FS}$ with a ratio of 10:1. The attack and release time is set to 10ms and 100ms respectively. From this information one can infer the behaviour of the system.

There should be three sections in the time domain plot: the first is when the AM signal is too low to trigger compression; the second is when bilateral compression occurs; and the third is when the steady-state signal gets attenuated to a magnitude that is unable to trigger compression.

If the level of the steady-state signal exceeds the threshold, but the level of the ramped AM signal is lower than the threshold, then the behaviour is unidirectional. In other words, the AM signal will be compressed and the steady-state signal will remain unaltered. The shape of the envelope for the AM signal will also be changed, adhering to the attack and release time that was set on the parameter control.

If the output from both tracks exceeds the threshold, then both the signals will be compressed. The steady-state sine wave will fluctuate following the frequency of the modulation on the AM signal. The AM signal will be kept at a moderate but gradually increasing magnitude. The shape of its envelope will still be altered.

When the output of the AM signal exceeds the threshold, but the steady-state sine is lower than the threshold, the behaviour will be unilateral whereby the AM signal will not be altered and the steady-state signal will undulate rhythmically, following the frequency of the modulation.

Again, these behaviours are expected. The AM signal is constantly compressed because compression is not triggered by itself, but by the steady-state sound. In other words, the level of the steady-state signal consistently exceeds the threshold, therefore the AM signal undergoes steady compression. The effect is akin to having the faders constantly kept at low levels. For the steady-state signal, the level fluctuates because compression is triggered by an AM signal.

3.4 Discussion – Bilateral Compression System

Results from the tests undertaken using sine waves and AM signals have shown that a bilateral compression system has been created successfully in Max 8. This was demonstrated by the system's unique characteristics as displayed on the time domain and frequency domain plots which are:

1) A compression curve that corresponds with the values of the two thresholds in the gain reduction algorithm.
2) The recursive nature of the system (i.e., the compression behaviour) is not only determined by the parameters set to the track, but also how the tracks interact with each other.
3) The occurrence of cross-modulation in the system demonstrated in the frequency magnitude plots.

These characteristics can then be used as indicators for future work to identify whether or not a system is built based on bilateral cross-adaptive architecture. Furthermore, using the algorithms provided in the previous chapter, the output from the system can be calculated[5] and compared to those produced by the DAWs in the previous experiment.

[5] if enough parameter details are provided.

Using the bilateral system developed in this section as the foundation, the process of expanding it to become a MIMO cross-adaptive architecture (i.e., nonlinear mixing system) is fairly straightforward. This is achieved by first sending the outputs from the tracks other than itself to a summing bus. The value from the bus is then used to calculate the amount of gain reduction. As was mentioned earlier, this is not ideal in the analogue domain because the number of connections and patch bay sockets needed to achieve this becomes impractical.

However, creating this architecture in Max 8 is not a difficult task because it only involves connecting the right outlets and inlets of Max objects which are, in this case, the ‘+’ symbol.

3.5 Nonlinear Mixing System

A nonlinear mixer based on the previously described bilateral compression system was created for this research. The architecture is the same as that system; the only difference is that it is expanded to include more tracks using summing busses.

The gen~ object for the nonlinear mixing system utilises 56 inputs and 24 outputs. This allows unique parameter settings for each track. Sixteen of the outputs correspond to the eight stereo tracks and the other eight outputs are for the outputs of the control signal for monitoring.

The mechanism for the architecture utilises busses to aggregate the output signal from all the tracks in the system (other than itself). The result is then used to calculate the control signal. The block diagram of the system was shown previously in Fig. 3.3. This creates a system that is inconstant and dynamic where magnitude oscillations on one track will affect the amount of compression on the others and in return will also affect itself. Encapsulating this architecture in one gen~ object allows for simultaneous sample-by-sample based processing. Therefore, the effect of one signal on the whole system is near-instantaneous.

The user interface for this system (Fig. 3.32) includes individual ‘Play’ buttons for each track to allow the audio loaded to the track to be played or stopped independently from one another. A ‘Start/Stop’ button that allows all the tracks to be played simultaneously is also included. Each track also includes an individual ‘Loop’ button allowing for independent cycling for each track.

There is also a sidechain enabler/disabler button on every track. When the sidechain is disabled, the track is sent via a different route that applies auto-adaptive compression and the audio signal is not included in the nonlinear mixing algorithm. Therefore, the user is able to control the number of tracks to be processed (a maximum of eight stereo tracks in this system).

The inclusion of the sidechain enabler/disabler button also simplifies operating the nonlinear system, especially for less experienced users. Consequently, rather than having to learn how to route the signals to achieve sidechain compression, the user may then produce the same effect just with the click of a button. Other than that, there is a system bypass toggle which provides a clean bypass of the whole system without altering the value of the parameters to enable quick A-B comparison.

Fig. 3.32 : *The user interface of the eight-track nonlinear mixing system.*

3.6 Multi-band Bilateral Compression

3.6.1 Introduction

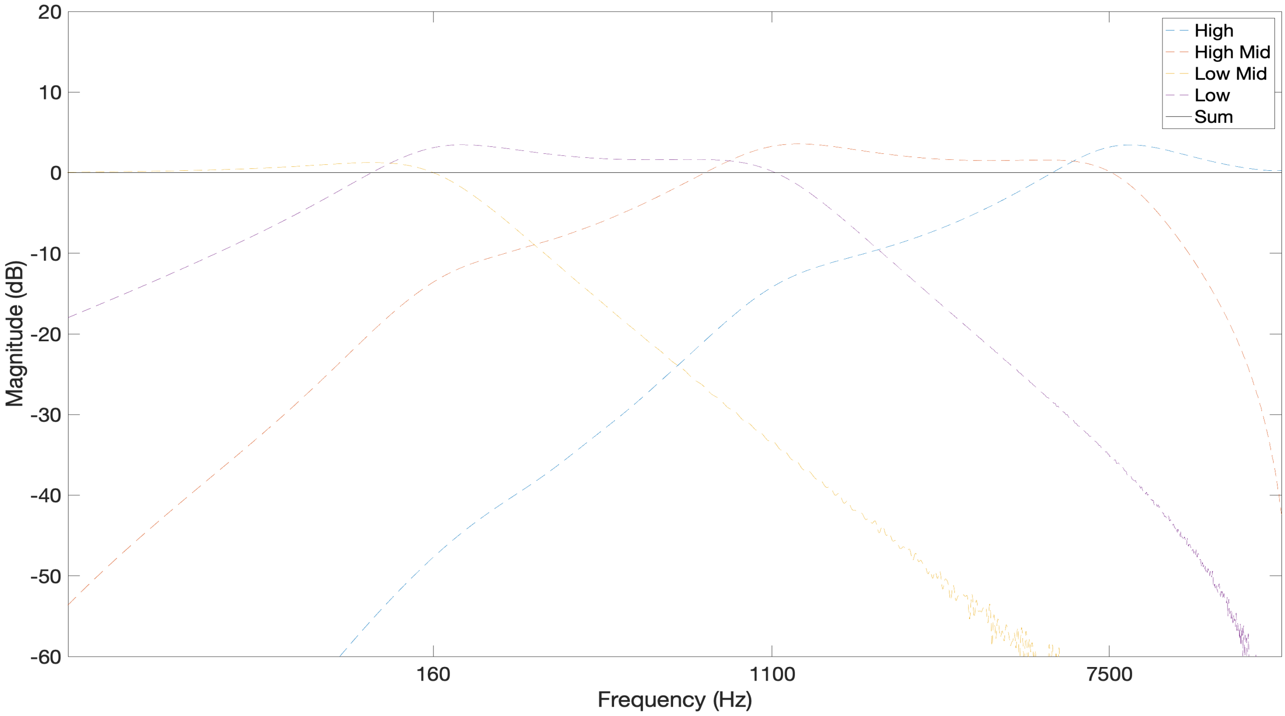
In this section, an alternative two-track multi-band compression architecture is proposed. This is achieved by dividing the spectral content of a track into four frequency bands. The novelty of this system is that each band is used to sidechain the others utilising the bilateral system. This method of implementing a bilateral compression algorithm is unique to this research and, to the best of the author's knowledge, has not previously been implemented.

To determine which band-splitting algorithm is most suitable for this purpose, the characteristics of the multi-band compressor in the aforementioned DAWs were investigated. This was conducted by measuring the impulse response of Reaper's ReaXcomp and LPX's Multipressor. The purpose was to examine the behaviour of the digital signal processor in both DAWs. It also serves as a benchmark for the band-splitter created for this research.

The impulse used for this experiment was generated by Max 8 using a function called 'click~'. First, it was simultaneously sent to all frequency bands before the result, including the summed output of the bands, is then finally plotted onto a frequency magnitude graph. For this test, the crossover was set to 160Hz, 1100Hz and 7500Hz, which are the default values in LPX's Multipressor.

3.6.1.1 Frequency Domain - DAW

The graph for both can be seen in Fig. 3.33. All of the four frequency bands in ReaXcomp exceeded the value of the summed output by about 5dB_{FS}. For Multipressor, however, there is a slight gain reduction in the second and third frequency band even though it generates a perfect reproduction of the signal. Additionally, for both signal processors the filter bands are not very well localised. For example, the lower mid frequencies for LPX and Reaper spans from 1Hz to about 21kHz. This is a trade-off for having a less steep slope because a less gradual filter slope will have a longer time domain response.



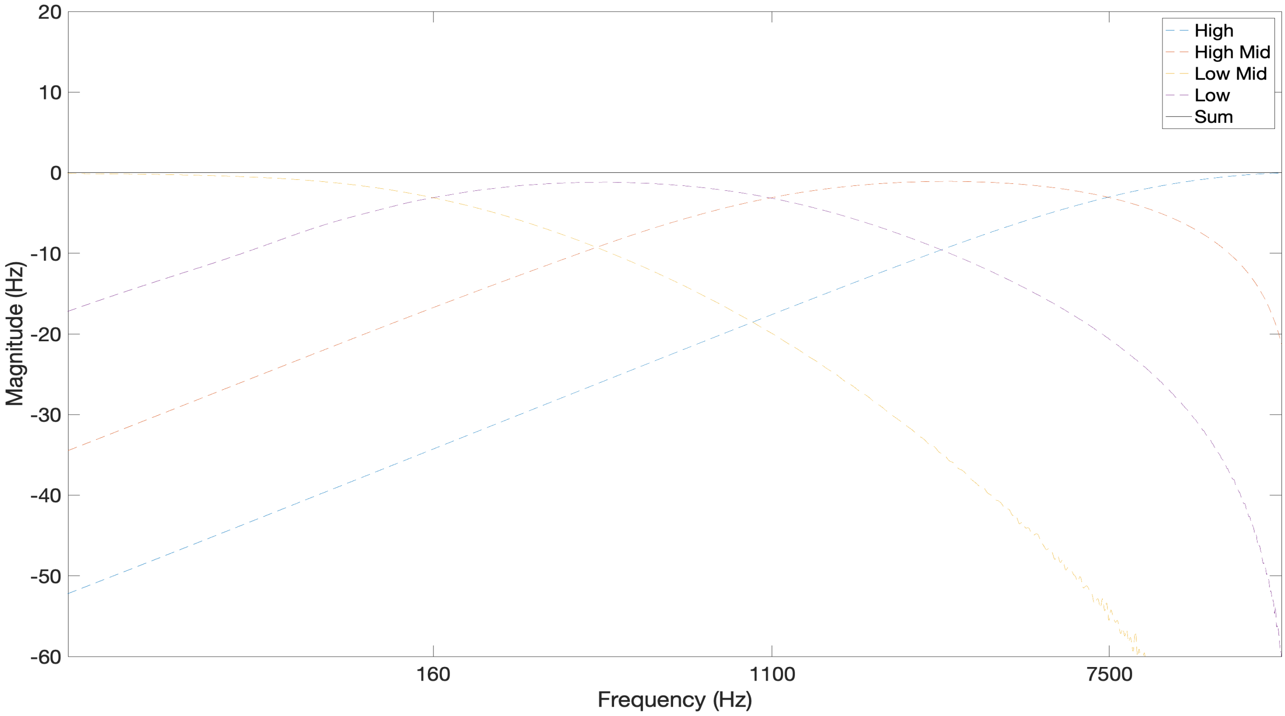


Fig. 3.33 : *Frequency magnitude plot of Reaper's ReaXcomp (top) and LPX's Multipressor (bottom)*.

Even though the summed impulse response for both band-splitting algorithm produced a perfect reconstruction of the initial input signal, the results from these architectures are not suitable for implementation in a nonlinear multi-band compression system. This is because the system uses the magnitude of each of the frequency bands to compress one another, therefore, a band-splitter algorithm that produces better localisation of each division is a better option. This is so that the content of one frequency band will not contribute to the calculation of gain reduction in another band. For example, in Fig. 3.33 the low-frequency band overlaps the high-frequency band which means any peaks on one band can affect the other. Furthermore, a more uniformed and predictable magnitude of each frequency band will lead to a better-informed decision-making process.

This research utilises Linkwitz-Riley crossover filters to split the audio into four frequency bands. The band filters are created by cascading two second-order Butterworth filters (Bohn 2005). This type of crossover, in theory, provides a flat amplitude response with no peaking at the point where the frequency response curves intersect. This is because there is a -6dB gain in the cutoff frequency for each band, resulting in a 0dB gain at the crossover.

Furthermore, there is no phase differences at the crossover due to the use of the low pass and high pass filters that are in phase. Implementing the LR-4 crossover in Max 8 is fairly straightforward with the use of the ‘filterdesign ’and ‘cascade ’objects. First the full spectrum of a track is divided into two halves; high and low. Each of those halves is then split into two parts resulting in a low band (band 1), low mid (band 2), high mid (band 3) and finally high (band 4).

3.6.1.2 Frequency Domain – Max 8

Using the same method as the previous section, the frequency magnitude plot of the system is produced to examine its behaviour. The result is shown in Fig. 3.34.

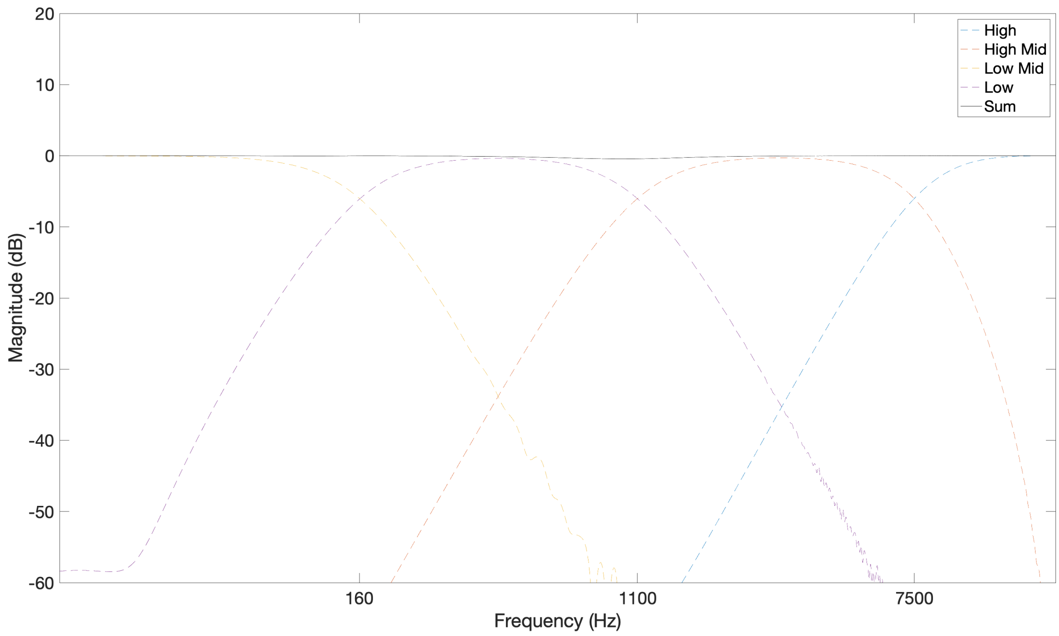


Fig. 3.34 : *Frequency magnitude plot of the Linkwitz-Riley crossover implemented in Max 8.*

The frequency magnitude plot for Linkwitz-Riley crossover implemented for this research displays steeper cutoff frequency slope. As a result, the frequency bands are more localised which enables better separation between each frequency bands. This feature is desirable for this system for the reason being that the low-frequency content will then not contribute to the amount of attenuation in the high-frequency region and vice versa.[6]

Furthermore, all of the bands produced identical peak magnitudes. Therefore, the behaviour is more predictable because there are no disparities between the maximum magnitude of a band and the maximum level of the summed output. Even though there is a slight deviation at the second crossover frequency, the benefit of using this architecture outweighs the drawbacks.

[6] Except for situations where cross-band modulation is intentional such as the one created for the final experiment, Page 139.

3.6.2 Routing

Now that the band splitter architecture has been established – the routing for a two-track nonlinear mixing system is explored. First, both the tracks are split into four frequency bands after which the system sidechains the same frequency bands bilaterally (Fig. 3.35). A thorough search of the relevant literature yielded no results pertaining to the use of such an audio processing architecture in music production.

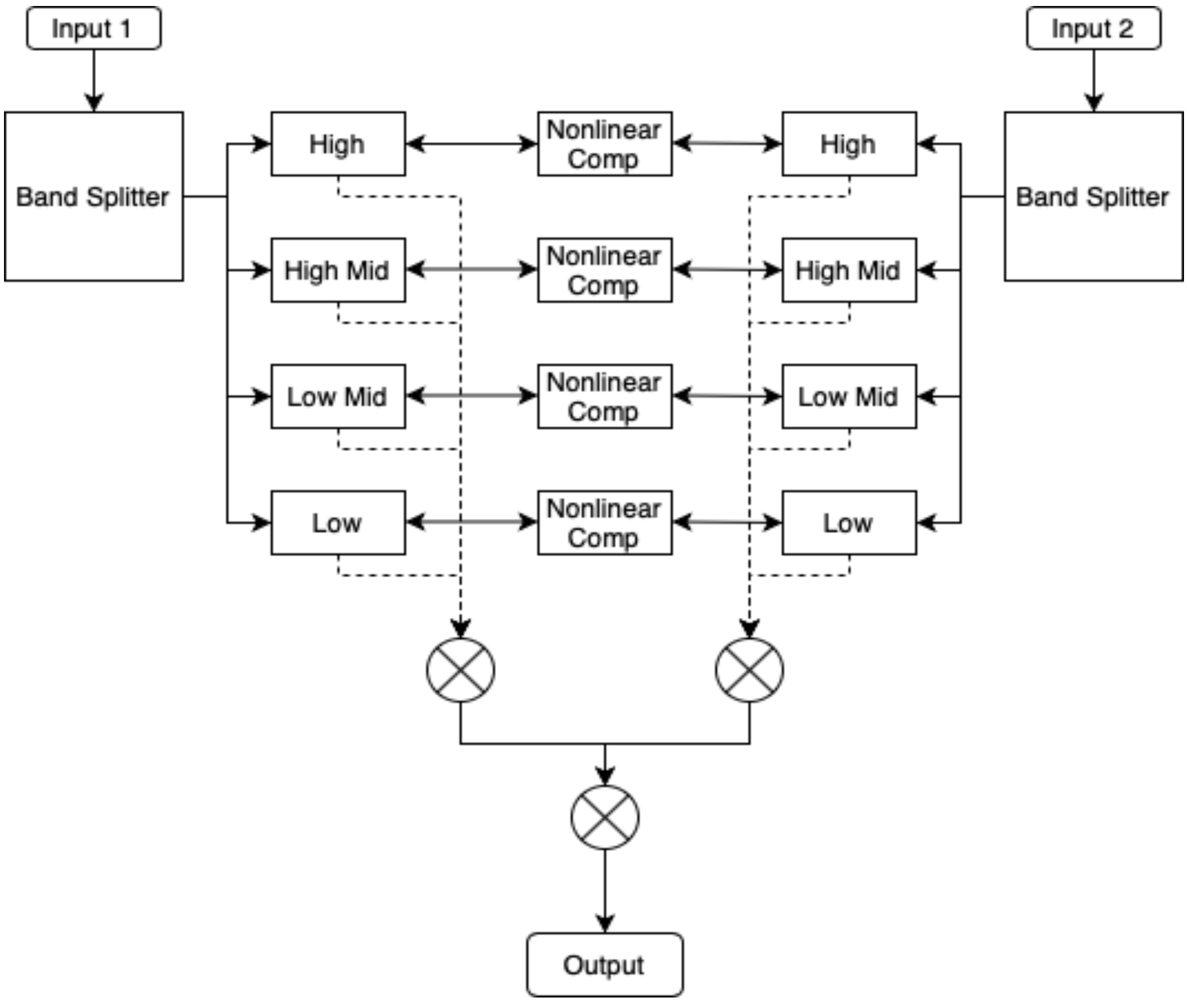


Fig. 3.35 : *Signal flowchart of a two-track nonlinear multi-band compressor.*

3.7 Conclusion – Architecture and Compression Characteristics

In this chapter, three novel compression architectures were explored. The first was a *bilateral compression system*, which served as the basis upon which to build the *nonlinear mixing system*. By implementing a fourth-order Linkwitz-Riley four-band splitter in the architecture, the bilateral cross-adaptive architecture was then expanded to become a *multi-band nonlinear mixing system*.

The bilateral compression system was first constructed in two DAWs; LPX and Reaper. This is because of the available user guidance mentioning the ability of both DAWs to implement feedback. Two tests were conducted to investigate whether the systems produce the expected behaviours; correlation to the parameter values and the occurrence of bilateral cross-modulation.

The first one used a dynamic and steady state signal to observe whether the gain reduction corresponded to the value of the two thresholds used to calculate gain reduction. The second test used two AM test signals to observe cross-modulation artefacts. The result from both experiments implies that neither DAW supports mutual cross-modulation in the way that the system proposed in this thesis requires. This implies that a bilateral compression system cannot be implemented in those workstations. This might be because the software uses sequence sensitive sample block processing.

The bilateral architecture was then constructed in Max 8. This programming environment was chosen because of its ability to facilitate feedback and also sample-based processing. The same tests conducted earlier were repeated using the newly developed system. The result is a compression curve that is consistent with the parameters set within the system.

Additionally, the result from the second test exhibits imprinting of cross-modulation artefacts from one signal onto another and vice versa. Both these reports verified the occurrence of cross-modulation within the system created in Max 8.

The bilateral compression system was then expanded to eight tracks. A further expansion of the architecture of this system included the use of fourth-order Linkwitz-Riley crossover to split two audio tracks into four bands. Each separate band was then processed using the bilateral architecture.

The next chapter provides three examples of work produced using the nonlinear mixing system and multi-band bilateral compression system; the first piece was mixed using moderate parameter settings, the second piece with parameters set to trigger overt amplitude modulation or pumping of the elements and the third piece mixed using the multi-band nonlinear mixing system. A detailed analysis of the musical examples is then provided using a myriad of analytical tools.

4 Musical Works and Analysis

4.1 Introduction

In general, the use of compression can be divided into three timeframes. In the beginning it was used as a protective tool to prevent damage to broadcasting equipment. Then it was used as a corrective tool to rectify inconsistent dynamics, or to de-ess a sibilant track, for example. Finally, it was utilised in production as a creative tool, to induce amplitude modulation, to shape the envelope of an audio signal or to enhance a reverb tail using parallel processing.

Similarly, mixing can also be grouped into two categories; corrective and creative. An insight into the approaches taken in mixing have highlighted different ideologies that audio engineers may use for their works. One is to produce the best faithful representation of the recording[7] and the other is the 'idiolect' approach or the engineer's interpretation of the recorded material (Marrington 2017)

One good example for the first approach is when a string ensemble and piano sound muddy when played together. The audio engineer may then use the equaliser to scoop out the low-mids of the strings to solve the problem. An example of the second approach is when an engineer uses an equaliser to make a piano sound like it is played back on a transistor radio so that it suits the stylistic direction for a section in a song.

This chapter explores a few scenarios where the nonlinear mixing system can be used. One option is to use the system in a subtle manner. In this situation, the signals will compress one another whereby the one with the higher level will be the dominant element and attenuate the others. The relationship between the signals for this approach vary over time, depending on their level. The parameters for this application are typically set to produce only slight compression. For example, threshold values that are slightly lower than the peak level with a low compression ratio.

[7] What is meant by faithful representation in this context is the representation of a musical audio in a way that does not attract attention to the processing which has been applied.

The other approach is to set the parameters to the extreme and to let the system process the audio independently. The difference between this usage of the system with the previously explained application is that the latter produces subtle ducking which may have potential in alleviating masking and providing a form of mix automation. By contrast, the output for this approach is meant to sound unique.

The following section provides three examples of how such a system can be utilised in a mixing process. The first example demonstrates the use of the system with moderate parameter settings. For the second example, the parameters were set to more extreme values. The final example utilises the multi-band bilateral compression system. The purpose of this chapter is to examine how massively mutual compression can affect a mix and how a mixing method using this system differs from a typical mixing approach in DAWs.

4.2 Ambient Music

Eight stereo tracks were used for this experiment which consisted of rendered audio tracks and software synthesisers. The tracks are listed in Table 4.1. The artificial sounds were synthesised using the LPX sound library (Apple Inc.) whereas the acoustic samples were recorded prior. The musicians were told to play with a loose tempo to stimulate micro-temporal variations. Synthesised sounds were then sequenced using LPX's MIDI sequencer, using the recorded instruments as a guide.

Then, each synthesised track was bounced to a wav file and no further processing was applied to them, except for the bass drum. Compression and slight clip distortion were applied to the track to give it more 'punch'[8] and to add spectral content. Fade-ins and fade-outs for each track were then applied using volume automation. This prevents sudden volume overload in the first few milliseconds before compression is triggered. This also ensures that all the elements will have the same maximum peak levels when the music starts which provides equal chance of each being the dominant signal. Finally, all the tracks were peak normalised and bounced separately.

[8] To give it more 'punch', the compressor was set to a relatively high ratio with a low threshold leaving the attack control to adjust the envelope of the onset.

Track	Instrument	Source	Processing	Character
1	Bass Drum	Logic Pro X (Kick 2 - Beat Machine)	• Compression to add punch • MIDI - Same velocity	• Transient • Non-reverberant • Low-frequency
2	Clap	Logic Pro X (Orion)	• MIDI - Same velocity	• Short reverb • High-mid frequency
3	Gendang Rebana	Recorded	• Double tracked • Panned	• Transient • Non-reverberant • Short resonance • Full spectrum
4	Gong	Logic Pro X	• MIDI - Same velocity	• Long resonance • Low-frequency
5	Marimba	Logic Pro X	• Added Gamelan • Panned (36° left) • MIDI - Varying velocity	• Transient • Non-reverberant • Non-resonant • Mid frequency
6	Gamelan	Logic Pro X	• Panned (36° right) • MIDI - Same velocity	• Long resonance • Low-mid frequency
7	Electro-acoustic Guitar	Recorded	• Stereo spread	• Plucked • Resonant • Full spectrum
8	Ambient Guitar	Recorded	• None	• Very reverberant • High-mid frequency

Table 4.1 : Audio Samples for the First Nonlinear Mix Test.

4.2.1 Track Details

The first track is the bass drum that uses the sound ‘Kick 2 - Beat Machine’. The signal is not reverberant and is active in the low to low-mid frequencies (between 100Hz to 500Hz). The loudness level is fairly constant, and the length of each signal is identical.

The second track is the clap which uses the sound ‘Clap - Orion’. The signal is reverberant and the length of the audio including the reverb tail is about 0.8 seconds. This sample is active in the mid frequency range, between 500Hz to about 1.2kHz and its loudness level is also fairly constant.

The gendang rebana in the third track is a Malay traditional percussion instrument played by striking the leather drum head by hand. It was recorded twice, double tracked and panned 36° to the left and right following the cosine panning law. It has a unique playing style whereby different technique used to strike it will produce a different pitch and envelope. As a general rule, striking it at about a three quarter distance from the centre of the head will produce a resonant low note. Striking the head nearer to the edge can produce two types of tone, depending on the technique used. One is a short, percussive, slap-like timbre consisting of high-frequency content and the other a resonant or ringing tone consisting of mid-range frequency at about 500Hz.

The gong is the fourth track, which uses the sound from LPX ‘Indonesian Gamelan Gongs’. It plays two bars per note and provides the low-frequency content for the musical piece. The notes played by the Indonesian gamelan gongs are sustained for a long duration.

The marimba is the fifth track and is panned to 36° to the left. It plays broken chords in quavers and is slightly mixed with another sound called ‘Indonesian Gamelan’. The ‘Indonesian gamelan’ sound is panned to the centre and plays crotchets instead of quavers, one octave higher than the marimba. However, there is a quaver note delay in the signal chain for ‘Indonesian gamelan’, giving it the illusion of playing quaver notes instead of crochet. The delayed notes however are not as resonant as the source. In the sixth track, the gamelan plays the same register as the marimba.

It uses the same sound used in the third track called ‘Indonesian Gamelan Gongs’ and is panned 36° to the right to counter-balance the fifth track.

The electro-acoustic guitar in the seventh track plays broken chords and is panned to the centre. It was originally in mono, recorded using a DI box directly into an audio interface. Afterwards, it was first processed using an equaliser to remove frequencies below 50Hz. From mono it was then converted to stereo using an LPX native plug-in called stereo spread. This processor splits the high mid and high-frequencies into a selected number of bands and distributes the bands alternately to the left and right channel (Apple Inc.). The notes played on the electro-acoustic guitar are sustained.

On the final track is the ambient electric guitar. The guitar used an overdrive pedal before the signal was sent to a reverb processor to produce a very reverberant output.

Three tests were conducted using the tracks listed above. The first used the nonlinear mixing system with the compression engaged, set to the lowest threshold value (-65dB_{FS}) and highest ratio (20:1). These parameters should generate results that exhibit excessive modulation to facilitate comparison with the other test results.

The second test used the nonlinear mixing system but with the compression disengaged. For this test, the level of each track was balanced to closely resemble the first experiment and no further processing was applied to the tracks. This is the control.

The third test applies compression on all the elements using the bass drum as the trigger, mimicking the technique made popular by EDM. This serves as a comparison of both methods.

The results were then loudness normalised following the EBU R128 standard using MATLAB. Normalisation is done to prevent biased preference for louder mix.

4.2.2 Discussion – Extreme Parameters for the Nonlinear Mix

In this section, the character of the processed audio samples is examined. The music consists of three elements that have long resonance or reverberance and the rest are transient elements. The output for the linear and the nonlinear mix are predictable.

For the linear mix, the timbre for the elements is unaltered and it is straightforward to produce a satisfactory mix (if the end result is only to produce a mix where all the elements are statically balanced with regard to their magnitude).

When it comes to the nonlinear mix, there are a few things to be taken into consideration before setting the parameters of the compressors. It was established that for this test, the threshold and ratio were purposely set to extreme values to strongly exercise compression in the mix. This leaves only a few parameters available to shape a mix and balance the level of each track. One example is the release time that could be made shorter to produce a more dynamic mix. The input gain can also be altered to increase or decrease the effect of compression to a signal.

Provided with enough information, the effect of the nonlinear system on the audio signals can be inferred. The most obvious outcome is that considering all the signals were normalised beforehand, all the tracks will be compressed by a significant amount because of the extreme ratio and threshold.

Furthermore, the envelope for every signal may be altered. This is because compression is caused by other elements in the mix. Therefore, it is possible that gain reduction has already been triggered even before the onset of a signal. This will affect the envelope as the duration of the rise will be made longer. The tail for the signals with a long decay might fluctuate, following the rhythms of the other elements. Signals with a short decay might be made even shorter.

For the signals with a constant dynamic level (for this experiment; bass drum, clap, gong and gamelan), there will be fluctuations in the dynamics of the output. Consequently, the overall dynamic balance of the mix will be constantly changed compared to the linear version. For example, it may well be that the bass drum is prominent in one section and in the next section it might be pushed back by the gong. As a result, there is constant movement and change within the mix. Complex interactions between different instruments are present which create interest for the listener.

Hence, careful decisions must be made while setting up the values of the parameters. It is not as straightforward as the use of compressors in an auto-adaptive architecture. The difference is clear as the values of the parameters for the auto-adaptive compressor will only affect its own track, and not another. For example, a compressor with the parameters set to a slow attack and fast release will make a signal punchier (as mentioned earlier). However, for a nonlinear system it will produce a different result altogether.

4.2.3 Nonlinear Mix (AudioFile1.wav)

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	0dB	-65dB$_{FS}$	20:1	5ms	150ms	20dB
Clap	+7dB	-65dB$_{FS}$	20:1	5ms	250ms	20dB
Gendang	+3dB	-65dB$_{FS}$	20:1	5ms	500ms	20dB
Gong	-2dB	-65dB$_{FS}$	20:1	5ms	500ms	20dB
Marimba	-1dB	-65dB$_{FS}$	20:1	5ms	50ms	20dB
Gamelan	0dB	-65dB$_{FS}$	20:1	5ms	50ms	20dB
El-Ac Guitar	-4dB	-65dB$_{FS}$	20:1	5ms	500ms	20dB
Amb Guitar	-1dB	-65dB$_{FS}$	20:1	5ms	500ms	20dB

Table 4.2 : Audio Samples and the Compression Parameters for the Extreme Nonlinear Mix.

The parameters were initially all set to the same value; 0dB input gain, -65dB$_{FS}$ threshold, 20:1 ratio, 5ms attack, 500ms release and 0dB make-up gain.

As a result, the peak level for all the tracks was significantly reduced to the values between -34dB$_{FS}$ to -27dB$_{FS}$. A make-up gain of 20dB was applied to all the tracks to compensate for the reduction in peak level. This adjustment does not interfere with the interaction of the tracks because the make-up gain is applied at the output stage, post compression. The gain compensation raised the peak level of the tracks to between -7dB$_{FS}$ to -15dB$_{FS}$.

Afterwards, the input gain for a few tracks was slightly altered to control the interaction between signals the in the system. Raising the input gain will make the track more dominant and less affected by compression. For example, the 'clap' is not prominent in the mix therefore the input gain was raised to allow it to modulate the amplitude of the other elements.

The input gain on the bass drum was not changed. The gong was too dominant in the mix therefore the gain was decreased. For the sustained track (electro-acoustic guitar and ambient guitar) the gain was reduced so that it was more affected by the transient signals.

Another important parameter that affects dominance in the system is the release time. As stated in the previous chapter, the track with a shorter release time will recover from compression more quickly than those with a longer release time, making it more likely to be the dominant track. Therefore, the release time for some of the tracks was adjusted to increase or decrease its dominance in the mix. The attack time for all the tracks was set to 5ms. A fast attack time will enhance the effect of compression in the system.

Having considered all the details, the resolved values of the parameters are shown in Table 4.2. The overall balance for the elements in the mix were obtained by the effect of compression, therefore the value of the make-up gain (which also functions as the faders in this system) was not changed.

4.2.4 Linear Mix (AudioFile2.wav)

As mentioned previously, the second mix only involves balancing the level of the audio tracks with no further processing. First, the compressors were disengaged by setting the threshold to $0dB_{FS}$ and ratio to 1:1. Then, by adjusting the input gain, the level of each of the tracks that was loaded into the nonlinear mixing system were adjusted aurally to closely resemble the result from the first mix. The process is straightforward as there were no other effects applied on the tracks. Thus, the timbre of the audio signals was not altered. The compressor settings are given in Table 4.3.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	-6dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Clap	-3dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Gendang	-10dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Gong	-9dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Marimba	-8dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Gamelan	-7dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
El-Ac Guitar	-14dB	$0dB_{FS}$	1:1	5ms	500ms	0dB
Amb Guitar	-12dB	$0dB_{FS}$	1:1	5ms	500ms	0dB

Table 4.3 : Audio Samples and the Compression Parameters for the Linear mix.

4.2.5 Sidechain Mix (AudioFile3.wav)

The third mix implements a multi-output cross-adaptive system in Max 8. This was achieved by making some modifications to the nonlinear mixing system. Instead of using the summed output of the other tracks to trigger compression, it is instead induced by the bass drum. This provides a similar effect of creating the sidechain compression architecture typically used in EDM as explained in the first chapter. The values of the compressor parameters are given in Table 4.4.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	0dB	$0dB_{FS}$	1:1	5ms	500ms	-15dB
Clap	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
Gendang	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
Gong	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
Marimba	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
Gamelan	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
El-Ac Guitar	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB
Amb Guitar	0dB	$-30dB_{FS}$	4:1	5ms	500ms	-10dB

Table 4.4 : Audio Samples and the Compression Parameters for the Sidechain Mix.

4.2.6 Commentary – Aural Analysis

Now that the different approaches to mixing the samples have been explored, this section provides a commentary on the music. Aural comparisons of the results have been conducted and have identified a few contrasts. The audio examples can be found in the accompanying materials that accompany this thesis.

One obvious difference is the amplitude modulation that occurred on the nonlinear and the sidechain mix. As expected, the amplitude did not fluctuate on the linear mix because there is no compression applied. That being said, listening to the nonlinear and the linear mix consecutively on a pair of headphones[9] revealed a few minute tonal differences, particularly in the low mid frequencies.

Further aural analysis of the nonlinear mix indicated that the gong and the bass drum are identical in level. This is obvious because both the tracks were programmed to have the same velocity. However, the level of the former seemed to be less prominent in the nonlinear mix, making the output sound wider and more spacious compared to the other mixes. This is caused by mutual

[9] The headphones that the author used to mix are a pair of AKG K92.

compression amongst the low-frequency elements in the music. When the bass drum and the gong are played simultaneously, the onset of the latter is attenuated by the drum, thus affecting its onset magnitude. Therefore, because the gong is attenuated during the first few milliseconds, it gave the impression of the instrument being placed a bit further from the listener even though the level is identical.

Furthermore, considering that the bass drum and the gong are the dominant elements in the centre of the stereo representation, any changes in magnitude may affect their balance. Because both of these low-frequency elements attenuated one another and the decay of the gong was affected by the other widely panned elements, the magnitude of this middle element was affected and gave the impression of a more spacious mix.

Consequently, the width of the sidechain mix seemed to fluctuate in that it is wider when the bass drum is absent in the mix and contracts to the middle when the bass drum is present. This is because when the amplitude of the other more widely panned elements is modulated by the bass drum, the overall width is affected and gives the illusion of a narrower stereo field.

There are micro-temporal variations in all three mixes because it was purposely elicited during recording. Timbral discrepancies, however, only occurred on the nonlinear and the sidechain mix caused by sidechain compression. The timbral variations on the sidechain mix are more predictable because they are controlled by one element that is rhythmically constant. In the nonlinear mix, however, the undulations seem more erratic. This is because gain reduction is not determined by only one element in the mix but by a mixture of all the other tracks. The bass drum, being the dominant signal in the sidechain mix, controls the gain for the other tracks. Considering the bass drum plays a constant rhythm, the undulations are therefore rhythmically consistent compared to the nonlinear mix. Furthermore, even the bass drum is affected by compression. Therefore, the amount of compression caused by it is inconsistent. Track by track analysis in the next section highlights this effect in more detail.

4.2.7 Track by Track Analysis

This section will provide a detailed analysis of the tracks mixed using the nonlinear system. This is to provide a better insight into how the architecture affected different audio samples.

4.2.7.1 Track 1 – Bass Drum (AudioFile4.wav)

As discussed in the previous section, by listening to the mixes it can be observed acoustically that the magnitude of the bass drum varied unlike the level in the linear and the sidechain mix. The amplitude against time plot (Fig. 4.1) demonstrates this. Fig. 4.1 exhibits the contrast in level and onset envelope. One difference is in the amplitude where, for the nonlinear mix, the sample was not only attenuated but this attenuation also fluctuates. Moreover, the attack envelope was slower than the original source as a result of compression triggered by the other elements in the mix.

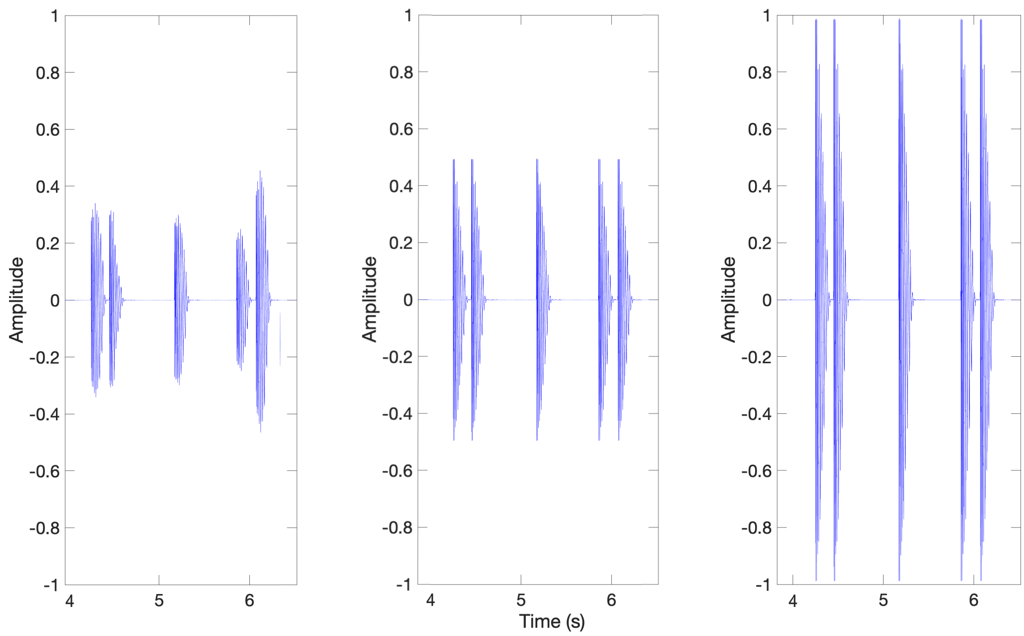


Fig. 4.1 : *Amplitude over time plot of the bass drum for the nonlinear mix (left), the linear mix (middle) and the sidechain mix (right).*

4.2.7.2 Track 2 – Clap

The reverb tail from the clap in the nonlinear mix is shorter compared to the other two mixes. In the sidechain mix, the decay was not affected because the clap occurs when the bass drum is not triggered, meaning that it is not affected by compression. As expected, the linear mix was not at all affected by the bass drum. The attack envelope for the nonlinear mix was not affected much. This is because the input gain was set much higher compared to the other tracks. The overall level, however, fluctuates. The amplitude variations on the reverb decay are very noticeable while listening to the soloed clap output of the nonlinear mix track.

4.2.7.3 Track 3 – Gendang Rebana

The original level for gendang rebana is fairly consistent throughout the mix and no obvious difference can be heard between the nonlinear and the linear mix. This is probably because the duration of the signal is too short for its envelope to be affected by compression in the nonlinear mix. Therefore, the other elements acted to keep the level of the signal constantly low.

For the sidechain mix, however, this track was heavily affected by compression triggered by the bass drum. Again, this is not a surprise as the gain attenuation affected all signals regardless of its level, unlike the nonlinear mix whereby the compression relied on the levels of all the other signals as well.

4.2.7.4 Track 4 – Gong (AudioFile5.wav)

The characteristics of the gong in the nonlinear mix are quite intriguing. In the aural analysis section, it was mentioned that the nonlinear mix sounded wider and the gong seemed to be less prominent in the music even though it gave the impression of having the same amplitude as the bass drum.

The time domain plot (Fig. 4.2) illustrates that the peak for the nonlinear mix does actually have the same peak amplitude as the linear mix. Also, the envelope of the audio signal is entirely different. The audio wave of the nonlinear mix is shaped by the other elements present in the track as was discussed earlier. Therefore, even though the peak level is comparable to the linear mix, temporal placement of the peak is different. This abnormality is inaudible in the full mix due to masking, which is why the gong seems to be less prominent in the nonlinear mix albeit sounding as if it has the same amplitude as the bass drum. However, listening to the solo output of this track, the level undulations can be heard clearly.

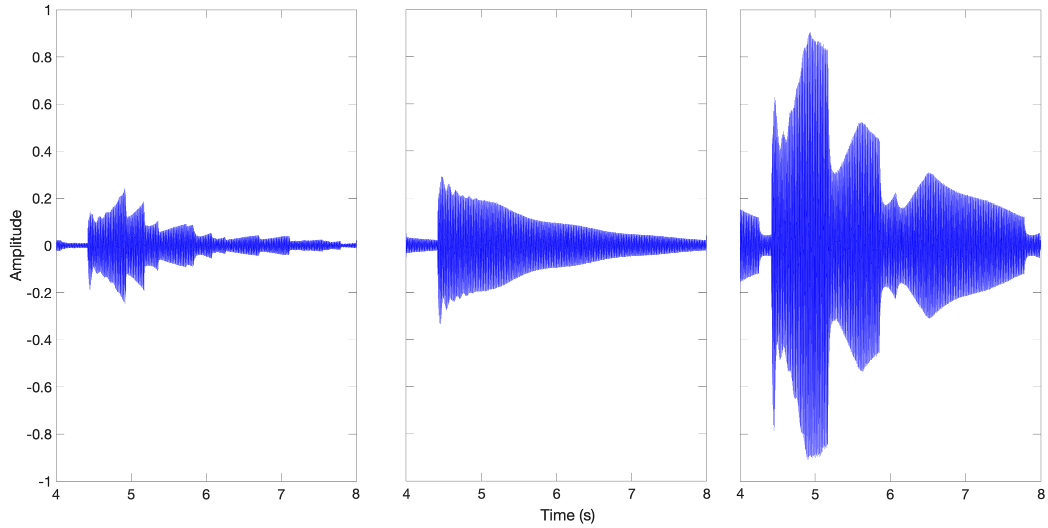


Fig. 4.2 : *Amplitude over time plot of the gong for the nonlinear mix (left), the linear mix (middle) and the sidechain mix (right).*

4.2.7.5 Track 5 – Marimba

Level fluctuations of the marimba were clearly audible in the summed output of the nonlinear mix. The track seems to fade in and fade out at times. The cause of this undulation is difficult to identify because there are too many variables. Additionally, the duration of the decay on the nonlinear mix is shorter compared to the other two mixes. Consequently, the timbre of this instrument sounds slightly unnatural because typically the marimba's wooden blocks will ring slightly after being struck. However, the ring is cut short because of the compression. This is even more obvious when the track is played solo. Furthermore, the envelope of this instrument changed. The attack is not instantaneous, unlike a typical percussion instrument.

4.2.7.6 Track 6 – Gamelan

The overall level of the gamelan in the nonlinear mix fluctuated even though it was programmed to a constant value. The same can be said about the resonant tail of the gamelan, whereas for the nonlinear mix the resonant tail is sometimes audible and sometimes it gets attenuated. Learning from the behaviours of the aforementioned tracks, this outcome is therefore predictable. Distinct amplitude oscillations can be heard in the solo output from the nonlinear mix. Similar to the gong, the decay of the resonance from the gamelan fluctuates in an abnormal manner.

4.2.7.7 Track 7 – Electro-Acoustic Guitar

In the nonlinear mix, the audibility of the guitar seems to fade in and out at times, similarly to the marimba. The ring from the notes played, however, are not cut short and can still be heard in the mix. This is probably because the magnitude of the sustain is strong enough to exceed the threshold and is continuous, resulting in this track having an effect comparable to a sustained element. On the solo track for the nonlinear mix, the amplitude modulation due to compression is clearly audible. However, just like the previous tracks, it is so complex as to sound random and it is difficult to identify the cause.

4.2.7.8 Track 8 – Ambient Guitar

For the summed output, auditive comparisons on the ambient guitar for the linear and the nonlinear mix do not highlight obvious differences. This is probably because of the nature of the audio sample itself. The timbre of the recorded track does give the impression of random pulsation and tonal variations. Therefore, when this signal is processed using the nonlinear system, the fluctuations are not so obvious considering timbral variations are within the character of the original signal. However, listening to the output in solo, the amplitude oscillations are clearly noticeable.

4.2.8 Discussion – Subtle Parameters for the Nonlinear Mix

In the previous experiment, the parameters for the nonlinear mixing system were purposely fixed to extreme values to heavily implement compression. The approach was to set the threshold to the lowest value and the ratio to the highest. The make-up gain was then applied to raise the level back to a more acceptable range. The input gain was only adjusted at the final step to increase or decrease the dominance of each element in the system.

The difference between the effects of compression on the nonlinear system and the sidechain system were easily identified. However, it was surmised that the audio example that was used may be unsuitable for extreme dynamic processing. Therefore, in the next example a more subtle version of the mix is produced. Using values that produce slight compression, the system may produce a better mix clarity without introducing any audible artefacts. Therefore, the use of compression in this experiment is corrective. Additionally, because compression is subtle and artefacts might be inaudible, a few of the analyses may be aided by visual representations of the summed output from the nonlinear system.

4.2.9 Nonlinear Mix - Subtle (AudioFile6.wav)

A different approach to mixing the tracks was taken compared to the previous experiment. For the previous mix the threshold and ratio were made constant and the mix shaped from that point onwards. For this method, however, the first step is different.

To begin with, the overall balance of the mix is first achieved by adjusting the values of the input gain. Once the desired balance has been achieved, the parameters of the compressor are then adjusted, starting with the threshold and ratio. The peak level indicator below the peak meter can be used as a guide to set the threshold. The aim is to induce slight amplitude modulation on the track. The final step is to use the make-up gain to compensate for the loss of level caused by compression.

Here, the amplitude modulation is only used to reduce the effect of masking in the mix. One example that can be heard is between the gong and the bass drum. In the linear mix, when they both play the downbeat at the start of a phrase the latter is masked by the former.

No adjustments of the attack and release time were needed because the mix was satisfactory using the values of 50ms and 500ms. The resulting parameter settings are given in Table 4.5. An analysis of the result from this experiment is provided in the next section.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	-6dB	-20dB$_{FS}$	6:1	5ms	500ms	0dB
Clap	-7dB	-12dB$_{FS}$	4:1	5ms	500ms	3dB
Gendang	-15dB	-20dB$_{FS}$	3:1	5ms	500ms	4dB
Gong	-20dB	-18dB$_{FS}$	3:1	5ms	500ms	9dB
Marimba	-14dB	-15dB$_{FS}$	5:1	5ms	500ms	5dB
Gamelan	-12dB	-15dB$_{FS}$	3:1	5ms	500ms	3dB
El-Ac Guitar	-19dB	-20dB$_{FS}$	3:1	5ms	500ms	5dB
Amb Guitar	-13dB	-13dB$_{FS}$	3:1	5ms	500ms	1dB

Table 4.5 : Audio Samples and the Compression Parameters for the Subtle Nonlinear Mix.

4.2.10 Commentary – Aural Analysis

For this section, the output from the system was analysed and compared with the result from the previous experiment. Through aural analysis, the difference between the nonlinear and linear mix is not particularly obvious except in terms of clarity. This is especially noticeable in the low-frequencies in the mix, exemplified by the gong and the bass drum in the spectrograms (Fig. 4.3). In the graph, the magnitude of the onset of the gong in the nonlinear mix (indicated by a purple box) is slightly reduced when the bass drum is present before it rises. This is not the case for the linear mix where the magnitude is high at the onset and fades over time.

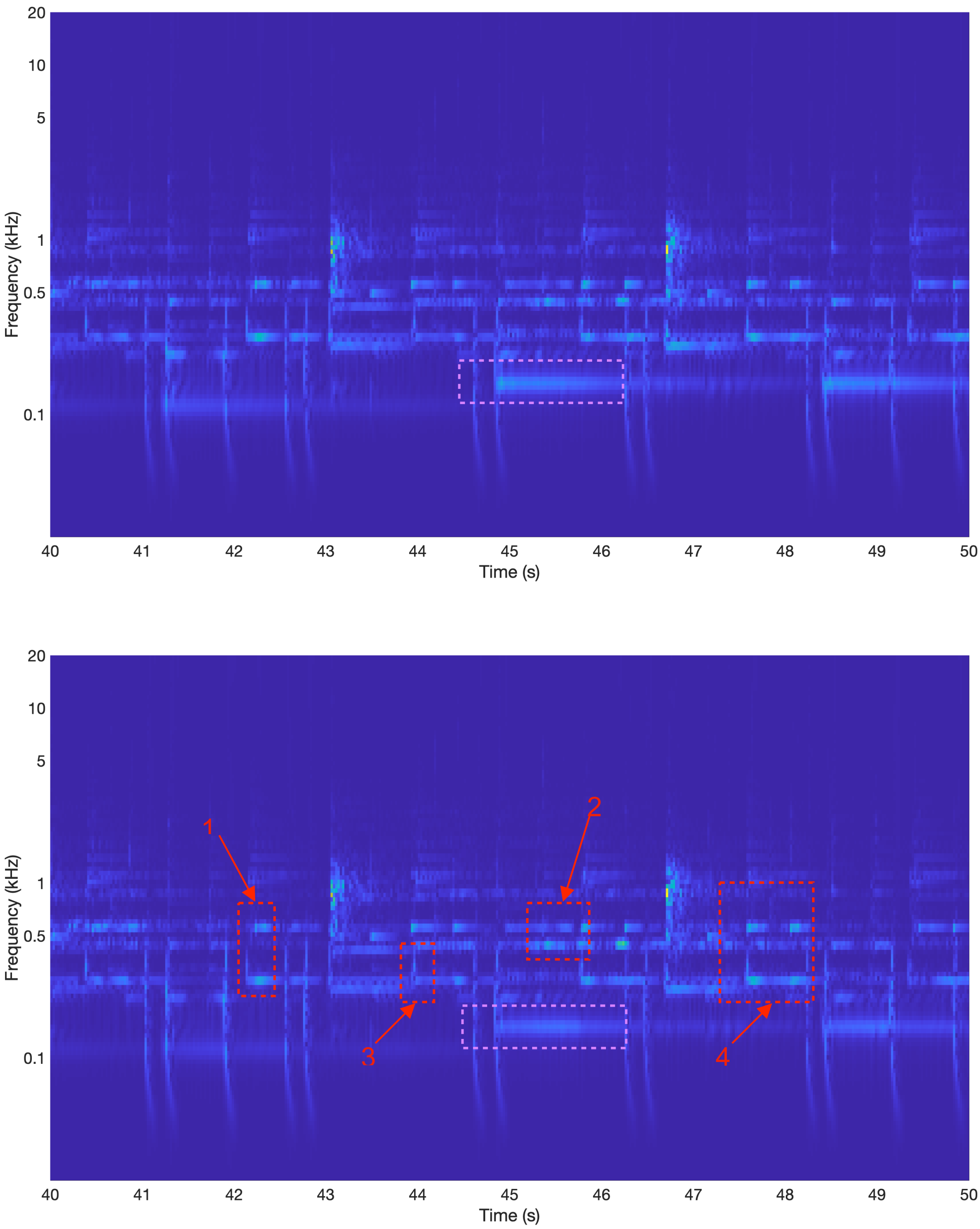


Fig. 4.3 : *Spectrogram of the linear mix (top) and the subtle nonlinear mix (bottom).*

4.2.11 Discussion – Other Analytical Tools

Although the spectrogram assisted in highlighting the differences between a linear and subtle nonlinear mix, the details are not apparent. Another tool that might be able to demonstrate how the compressor reacts to every element in the mix is the gain reduction against time plot, referred to in this thesis as the control signal.

4.2.11.1 Control Signal in Time Domain

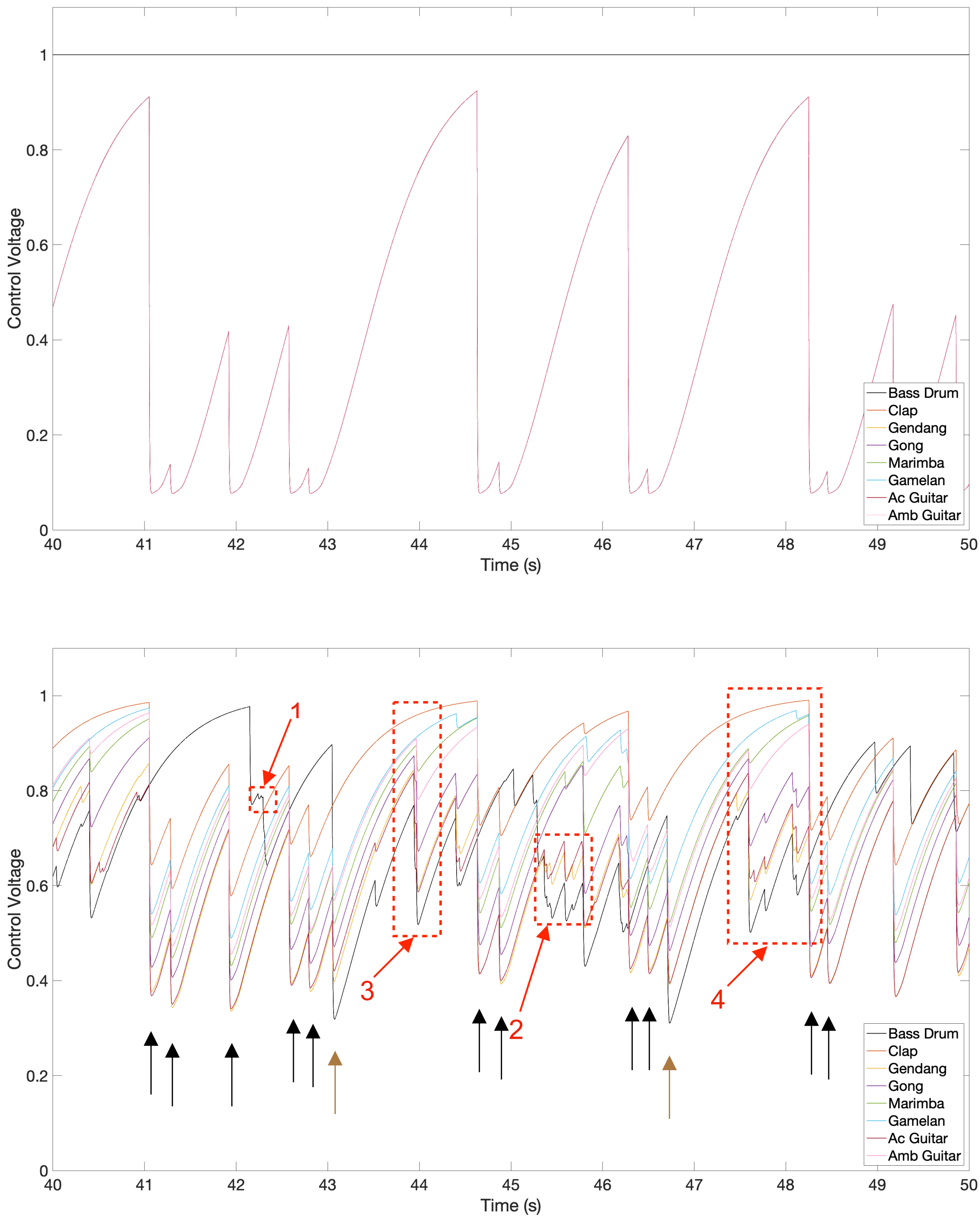


Fig. 4.4 : *Amplitude over time plot of the sidechain architecture's control signal (top) and subtle nonlinear system's control signal (bottom).*

Even though aural comparisons between the linear and the subtle nonlinear mix reveal minor differences, the time domain plot of the control signal shows that there is significant complexity. Fig. 4.4 shows stark contrast between a sidechain compression system with a nonlinear mixing system. As a guide, the value of 1 represents no compression and the value of 0 constitutes total gain reduction, or a complete silence. A linear system, for example, will have the value of the control signal at 1 for all tracks throughout the duration of the song.

Overall, the control signal plot corresponds with the parameters that were set. For example, a steep downwards line is consistent with the fast attack time of the compressor and the gradual rise is the release time. Additionally, by analysing the control signal plot of the nonlinear mix, the interaction between all the signals in the system can also be perceived.

For example, the red square marked 1 on the control signal plot pointed out sudden gain reduction on the bass drum. The spectrogram, presented prior, marked the presence of gong and marimba in that section. Therefore, it can be deduced that all of the tracks exceeded their respective threshold except for the bass drum. This is indicated by the direction of the curve where, while the curve of the drum was decreasing, the rest was rising. When the marimba and the gendang were present, they triggered compression to the bass drum. However, because the bass drum is a transient signal, the effect of compression is short-lived and inaudible.

Another example is the red square marked 2. Here, the level of the electro-acoustic guitar and the gendang did not exceed the threshold and thus the gain for both the tracks are reduced. The rapid oscillations on both control signals are most likely triggered by the marimba as shown in the spectrogram. This is because the marimba is the most active signal in that section and most certainly has exceeded the threshold value that was set to it. Minor fluctuations can also be observed at 0:47 caused by both the dominant gong and marimba.

In the box marked 3 on the control signal plot, all the tracks were compressed except for two which are the gamelan and the clap. Therefore, it can be deduced that even though the tracks were recovering from compression, the level did not exceed the threshold, except for the gamelan. Thus, when the gamelan was played it triggered compression to all the other tracks. The clap was not affected because it had the threshold and the ratio settings, enabling the track to have a higher level overall.

One final example is in the box marked 4. Four tracks were compressed while four others were dominant. From what has been described previously, an explanation for this compression curve is quite straightforward. The rapid oscillations on the compressed signals are caused by the dominant transient signals (marimba and gamelan). The overall level is suppressed by the steady-state signal (ambient guitar). The most obvious gain reduction is triggered by the bass drum and the clap. The lowest points in the plot correspond to the occurrence of both elements as indicated by the black arrow (bass drum) and brown arrow (clap). The sidechain system in this experiment used the same parameter values for all the compressors, thus the control signal is identical. For the next test, different values are used.

4.2.12 Discussion – Potential Application of the Nonlinear System

The control signal plot demonstrates that the nonlinear mixing system can be used to reduce masking. It is worth mentioning that the marimba, electro-acoustic guitar and the gamelan have considerable overlap in the frequency domain, as do the gong and the bass drum in the low-frequencies. Using this approach, the audio signals within the same domain can be made clearer.

This behaviour is certainly different from the sidechain system, as shown by the control signal plot. The bass drum is usually the dominant signal in such architectures and triggers compression for all the tracks that are subjugated. There are no cross-modulations or any interactions between the tracks in the system and sidechain only occurs in one direction, meaning only the bass drum is controlling the other elements.

The technique of using sidechain compression to reduce masking is not uncommon. One of the more popular examples for the use of this approach is by using it to sidechain the bass to the bass drum. A thorough search of the literature, however, shows that the use of a massively mutual sidechain compression system to reduce masking has never been attempted before. It is therefore one of the novelties that is introduced in this thesis.

4.3 Experimental/Hip-Hop

As was previously mentioned, the audio examples used for the first experiment may be unsuitable for extreme compression. Therefore, for the next experiment a more appropriate audio example which emulates the style of Flying Lotus is used to showcase the possibility of the nonlinear mixing system to be used in an extreme setting.

Track	Instrument	Source	Processing	Character
1	Bass Drum	Logic Pro X (Simmons SDS-V)	• MIDI - Same velocity	• Transient • Non-reverberant • Low-frequency
2	Snare	Logic Pro X (Simmons SDS-V)	• MIDI - Same velocity	• Transient • High-mid frequency
3	Hi-Hat	Logic Pro X (Simmons SDS-V)	• MIDI - Same velocity	• Transient • Non-reverberant • High-frequency • Syncopated rhythm
4	Synthesiser Pad	Logic Pro X (Black Sun)	• MIDI - Same velocity	• Sustained • Low-mid frequency
5	Vocal Sample and Cowbell	Logic Pro X (Sample Library)	• Reverb (ChromaVerb)	• Transient • Short reverb
6	Percussion Sample	Logic Pro X (Solid Ice Percussion)	• None	• Transient • Non-reverberant • Syncopated rhythm
7	Percussion Sample	Logic Pro X (Native American Drum 2)	• None	• Transient • Constant quavers • Low-frequency

Table 4.6 : Audio Samples for the Second Nonlinear Mix Test.

4.3.1 Track Details

Here, seven stereo tracks were used which consist of sample loops and synthesised sounds generated using LPX. Similar to the previous test, all the tracks were first bounced down and normalised before fade-in and fade-out were applied, without additional processing except for Track 5. Reverberation was applied on the track to extend the decay. All the MIDI sequenced tracks have a uniform velocity of 80. Therefore, any level of undulations in the mix were imparted by the system. The instruments track list is given in Table 4.6.

4.3.1 Nonlinear Mix (AudioFile7.wav)

For this mix, the same approach was taken as the first example in the previous section. Firstly, the threshold and the ratio were set to the extreme values of -65dB_{FS} and 20:1. Using these settings, the level for the tracks were reduced to between -35dB_{FS} to -25dB_{FS}. Track 4, 6 and 7 were heavily affected by compression. Therefore, make-up gain was applied to it to counter this and to balance the level with the other tracks. To strongly exercise compression, the input gain on the synth pad and the percussion loop were decreased by 5dB and 10dB respectively.

To enable Track 7 to recover from compression, the release time was shortened (300ms). The input gain on the track was then decreased to compress it even further. Because the bass drum was too dominant in the mix, the release time was made longer to reduce its dominance (1000ms). The hi-hat was made more dynamic by lowering the input gain by 10dB and compensating the gain reduction by adjusting the make-up gain. The snare was also made more prominent in the mix by increasing the make-up gain.

As a result, the mix has a slightly similar amplitude modulated behaviour to a typical sidechain mix. The difference is in the characteristics of the compression, whereby for a sidechain mix the rhythmic undulation is consistent with the presence of the bass drum. However, for the nonlinear system, the gain reductions not only correlate with the bass drum but also with the other elements as well. As a final note, the bass drum is also affected by compression.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	4.5dB	-65dB_{FS}	20:1	5ms	1000ms	0dB
Snare	10dB	-65dB_{FS}	20:1	5ms	500ms	7dB
Hi-hat	-10dB	-65dB_{FS}	20:1	5ms	500ms	20dB
Synth Pad	-5dB	-65dB_{FS}	20:1	5ms	500ms	16dB
Samples	-6dB	-65dB_{FS}	20:1	5ms	500ms	13dB
Percussion	-10dB	-65dB_{FS}	20:1	5ms	500ms	20dB
NAD	-4dB	-65dB_{FS}	20:1	5ms	300ms	20dB

Table 4.7 : Audio Samples and the Compression Parameters for the Nonlinear Mix.

4.3.2 Linear Mix (AudioFile8.wav)

For the linear mix, the aim was to reproduce a balance of the audio samples similar to the nonlinear mix for comparison. No further processing was applied to the tracks. The parameters are shown in Table 4.8.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	0dB	$0dB_{FS}$	1:1	5ms	500ms	-6dB
Snare	0dB	$0dB_{FS}$	1:1	5ms	500ms	-10dB
Hi-hat	0dB	$0dB_{FS}$	1:1	5ms	500ms	-15dB
Synth Pad	0dB	$0dB_{FS}$	1:1	5ms	500ms	-7dB
Samples	0dB	$0dB_{FS}$	1:1	5ms	500ms	-14dB
Percussion	0dB	$0dB_{FS}$	1:1	5ms	500ms	-8dB
NAD	0dB	$0dB_{FS}$	1:1	5ms	500ms	0dB

Table 4.8 : Audio Samples and the Compression Parameters for the Linear Mix.

4.3.3 Sidechain Mix (AudioFile9.wav)

Similarly, for the sidechain mix, the goal was to reproduce a similar balance to the nonlinear mix. The bass drum was left unaltered as it is the dominant signal. For the other tracks, the compression parameters were individually set to produce a mix that exhibits overt compression that parallels the rhythm of the bass drum. Make-up gains were adjusted to balance the level of the tracks.

Track	Input Gain	Threshold	Ratio	Attack	Release	Make-up Gain
Bass Drum	0dB	0dBFS	1:1	5ms	500ms	-10dB
Snare	0dB	-10dBFS	4:1	5ms	570ms	-8dB
Hi-hat	0dB	-25dBFS	4:1	5ms	301ms	-5dB
Synth Pad	0dB	-31dBFS	8:1	5ms	99ms	-3dB
Samples	0dB	-40dBFS	9:1	5ms	100ms	-12dB
Percussion	0dB	-38dBFS	10:1	5ms	500ms	5dB
NAD	-6dB	-25dBFS	10:1	5ms	50ms	7.5dB

Table 4.9 : Audio Samples and the Compression Parameters for the Sidechain Mix.

4.3.5 Commentary – Aural Analysis

Aural comparisons between the sidechain and the nonlinear mix exhibited a few noticeable differences. The obvious one is in the bass drum in both examples. The bass drum in the sidechain mix is overpowering and dominates the whole song. In the nonlinear mix, although the bass drum triggers compression, it does not seem to do so for the entire duration. There are a few instances where the bass drum is slightly subdued by the other elements.

The spectral balance of the bass drum in the nonlinear mix also seems to change throughout the loop. This might be caused by masking whereby when the bass drum is lower in level, its presence is masked by other elements such as the shaker or the hi-hat. Pumping is apparent in both mixes. However, the magnitude of the rhythmic modulation in the nonlinear mix is not as intense as compared to the sidechain version.

Because the parameters that were set for the nonlinear mix were extreme, one might expect that the result would be more incoherent and disjointed. This is because only the dominant track will be audible, while the others get compressed to an inaudible level. However, this is not the case because all of the elements are pushing one another simultaneously, meaning that there is no completely dominant component in the system. Thus, the output level from the nonlinear system is very low. It was changed by the gain compensation which was applied during normalisation that raised its level relative to the other mixes.

One unexpected result from this experiment is the mixing approach. For the sidechain and the linear mix, the process of balancing the level of the tracks is quite straight-forward. It begins with the audio engineer choosing one track as a reference level (usually either drums or vocals) and adjusting the level of the other tracks accordingly, relative to the reference. They then build the mix from there onward, comparing the level one track after another. This method is also known as bottom-up mixing[10].

For the nonlinear system, however, the mixing will usually use the top-down approach. The audio engineer will have to listen to the mix as a whole, rather than listening per track. This is because the interaction between the tracks should be taken into consideration. For example, during mixing an engineer might want to make sure that one element is not too buried or too prominent. However, altering the level of one element in the mix might affect the overall balance. This situation demonstrates that the balancing process for the nonlinear system is multidimensional instead of one-at-a-time.

4.3.5.1 Control Signal in Time Domain

Unsurprisingly, the control signal plot for the nonlinear system is different from the sidechain mix. Because the parameter settings on each track in the sidechain mix were set to different values, the compression curve for each element is different. However, the rhythm of the amplitude modulation is still subservient to the bass drum. This is different from the nonlinear mix because the curve is erratic. In fact, even the bass drum is modulated because of the bilateral nature of the system. This is different compared to the sidechain mix where the bass drum is constantly at the value of 1, indicating no compression at all. Also, as mentioned in the previous section, the level of the output from the nonlinear system was low, as shown on the values of the control signal (Y-axis). This is because the gain reduction for all the tracks in the nonlinear architecture is excessive (Fig. 4.5).

---

[10] Not to be confused with the same term used by Phil Harding in a chapter he wrote for the book "Mixing Music - Perspectives on Music Production" (ed. Hepworth-Sawyer and Hodgson; Routledge, 2017). For this thesis, the term 'bottom-up' and 'top-down' follows the context that is used in a few online articles such as "5 Tips for Top-Down Mixing Effectively" (https://www.waves.com/tips-for-top-down-mixing-effectively), "Is a Top-Down Approach to Mixing Always Best?" (https://www.soundonsound.com/sound-advice/q-top-down-approach-mixing-always-best) or "Better Mixes with Top-Down Mixing" (https://www.sonarworks.com/blog/learn/better-mixes-with-top-down-mixing/).

One similarity between the control signal plot for both mixes is the position of the minimum points that coincide with the occurrence of the bass drum. This is because the bass drum is purposely set up to trigger compression in both examples. However, the values of these lowest points in the nonlinear mix change continuously unlike the sidechain. An interesting feature in the control signal plot of the nonlinear mix is the peak. All the peaks of the tracks did not overlap one another. This behaviour is caused by the extreme values that were set on the parameters. Because of this, as soon as one element exceeded the threshold, it immediately compressed all the other tracks excessively.

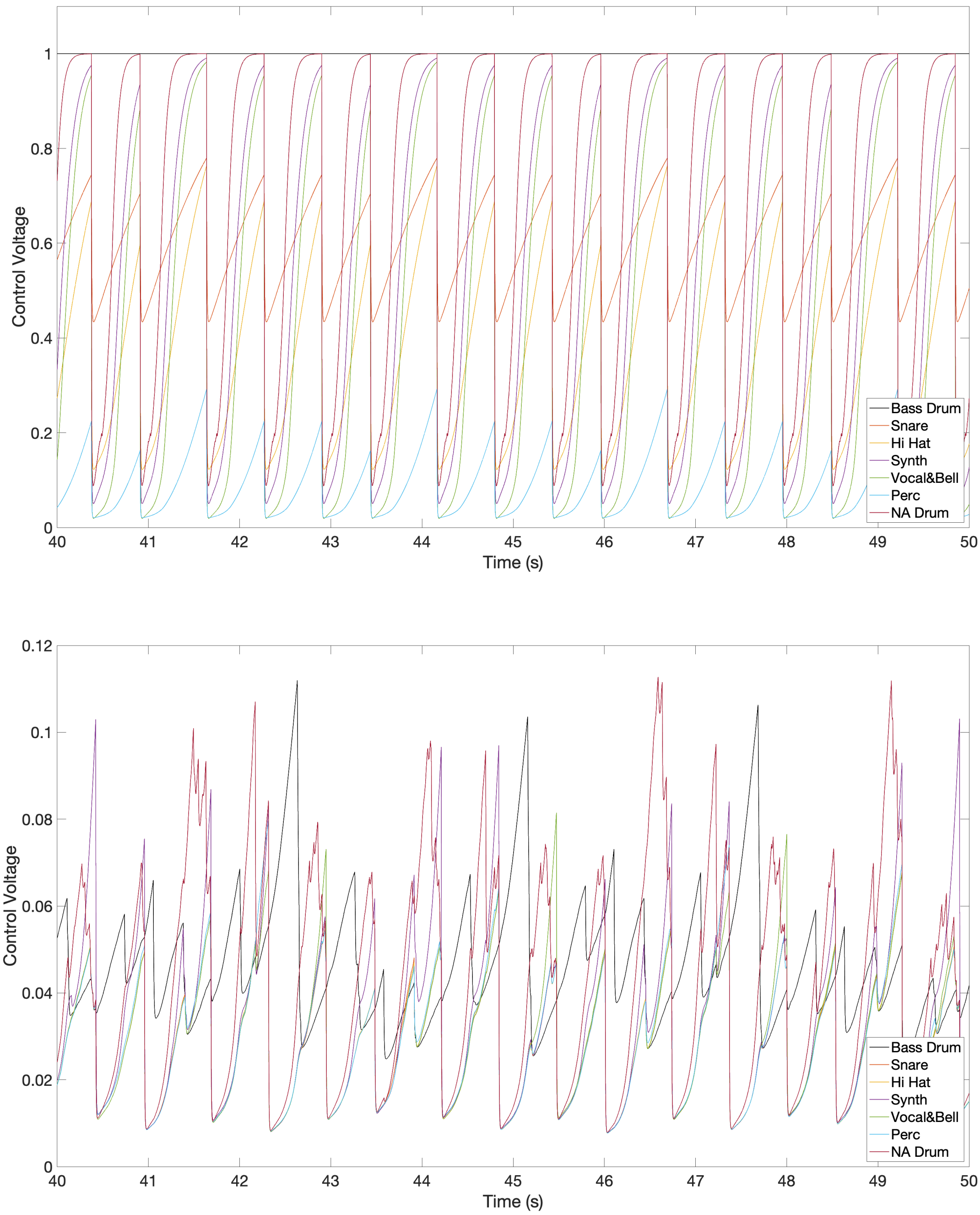


Fig. 4.5 : *Amplitude over time plot of the control signal for the sidechain mix (top) and the nonlinear mix (bottom).*

4.3.6 Conclusion – Nonlinear System

For the nonlinear system, compression of various intensities can be used to process different types of audio signal and provide heterogeneous results depending on the objective of the mix. In the examples given earlier, two processing approaches have been taken: subtle compression and overt compression. Subtle compression proved to be suitable to help to reduce masking. For extreme compression, the nonlinear mixing system is instrumental in shaping the mix and can be used creatively.

Analyses of the compression behaviour have been conducted using spectrograms. For transient signals and subtle parameters, however, the effect of compression might be inaudible. Therefore, the analyses were facilitated using the time domain plot of the control signal.

The control signal plot of the sidechain mix exhibits obvious differences compared to the nonlinear mix. One of these differences is the amplitude modulation. For the sidechain mix, it happens only when the dominant signal exceeds its threshold. For the nonlinear architecture, any signals that exceed the threshold will trigger compression for the tracks in the system.

The other difference was the mixing approach. Using the nonlinear system, the top-down approach was the more intuitive method because the user needed to consider the interaction between the tracks. For the sidechain mix, the user may take the bottom-up approach because the interaction only goes one way between the dominant signal and the subjugated tracks.

So far, this thesis has focused on the common method of applying a compression architecture where full frequency signals are used to trigger compression. In the next section, the use of another novel architecture will be explored and analysed. Instead of using the full spectrum of a signal to trigger compression on another, it uses determined frequency bands to reduce the gain of entire signals or any frequency bands that the user desires.

4.3.7 Multi-Band Nonlinear System

In the second chapter, another novel architecture was proposed which is the multi-band nonlinear system. The next section explores its application for mixing purposes, especially by using extreme parameters. The aim is to discover the characteristics of compression and how it affects audio signals.

To assist comparison with the nonlinear system, the same audio samples as the previous mix were used. However, they were grouped into two stems. In the first stem is the drums (hi-hat, snare, bass drum) and in the second, all the other samples (synthesiser pad and samples).

4.3.7.1 Mix 1 – Same-Band Compression (AudioFile10.wav)

For this experiment, the multi-band nonlinear system as illustrated in Fig. 3.35 was utilised. The first step was to pre-mix the samples by balancing the level of each track in the stems. This was done using the nonlinear mixing system with the compressor disengaged. Then, both stems were bounced and loaded into the multi-band nonlinear system.

On the system, the crossover frequencies of each of the stems were first adjusted to the values that best separated the content of each of the bands as indicated in Table 4.10. This was conducted aurally. After that, the parameters of the compressor for each frequency band on both stems were adjusted with the goal of inducing pumping. Finally, make-up gain was applied to compensate for level reduction. The result was bounced and then normalised in MATLAB.

Due to the novelty of this architecture, there is no perfect way to predict the end result. Therefore, the approach taken while mixing this system was taken to produce an outcome that enhances the quality of the mix with pumping as the main feature.

Frequency Band	Stem 1 (Drums)	Stem 2 (Samples)
Band 1 - Low	Bass Drum Body and Snare Body (0Hz - 250Hz)	NAD Body (0Hz - 200Hz)
Band 2 - Low Mid	Snare Attack & Bass Drum Attack (250Hz - 4.2kHz)	Ambient Pad, NAD Attack, 'Clavé' & Vocal (200Hz - 5.1kHz)
Band 3 - High Mid	Hi-Hat Presence & Bass Drum (distortion) (4.2kHz - 10.5kHz)	Shaker (5.1kHz - 10.3kHz)
Band 4 - High	Hi-Hat Brilliance (10.5kHz - 20kHz)	Brilliance and Noise (10.5kHz - 20kHz)

Table 4.10 : The Range of the Frequency and the Content in Bands 1 to 4 for the Multi-Band Nonlinear Compressor Test.

4.3.7.2 Commentary – Aural Analysis

By using this approach, the result was different from the results produced by the nonlinear mixing system. One obvious difference is that the pumping occurred only in the low-frequencies. It can be heard on the interaction between the native American drum (henceforth referred to as NAD) and the bass drum. In the original sample, the NAD plays quavers with a constant loudness level. However, the occurrence of the bass drum pushed the NAD out of the mix, giving the illusion that the NAD only played the upbeat. The interaction between the two signals also made the tone of the NAD more like a heartbeat beating irregularly.

Additionally, the tone of the bass drum is altered compared to the original tone in the stem. The presence caused by mild distortion in the sample used is inaudible caused by compression trigger in Stem 2. As a result, even though the bass frequencies in Stem 1 are dominant in a few instances, the impact is not as powerful for the nonlinear or sidechain mix. Furthermore, the high-frequency noise, which is an element of the ambient pad, is more evident in the multi-band mix. This is because when compression is triggered in the nonlinear and linear mix, the whole spectrum of the ambient pad is pushed down. For this system, only the relevant frequency bands are affected. Since Stem 2 is more active in Band 3 and 4 of this system, it constantly dominates the high-frequency spectrum making the noise more prevalent.

The low-mids of this composition are occupied by three elements which are the ambient pad's low-frequency content, the bass drum's attack and the snare. Pumping can be heard on the ambient pad triggered by the bass drum and the snare. Compression also occurs on Stem 1's low-mids because when a steady-state sound is compressing a transient signal, the effect is not audible.

This system produced a texture that is different from the nonlinear or the sidechain mix. When using this architecture, the pumping is not overt. Even though pumping can be heard in a few elements such as the NAD and the ambient pad, the whole mix sounds lifeless and dull. Another difference is in the behaviour of the high-frequency content, particularly the noise on the pad. As was mentioned earlier, the content of the fourth band of the second stem were barely modulated. The same can be said for Stem 1's hi-hat where the amplitude did not change throughout the mix.

Therefore, it may be that the attributes of the high-frequency content are essential for a mix to sound dynamic and for pumping to be easily detected. Thus, another type of multi-band nonlinear architecture is used in a mixing process to investigate this notion.

4.3.7.3 Mix 2 – Cross-Band Compression (AudioFile11.wav)

Since mutual compression only happens to the elements in the same frequency band, pumping is not as obvious as compared to the other experiments. Therefore, another architecture was created that enables low-frequency content from one track to compress the high-frequency bands of another and vice versa.

Referring to Fig. 3.35, the configuration for the multi-band nonlinear architecture was changed. For this experiment, the low-mids were utilised to compress the high-mids, the lows to compress the highs and vice versa. Therefore, the high-frequencies may be more affected by gain reduction. This is because the content of the low-frequencies is less active rhythmically, especially in EDM whereby it usually only plays 'four-on-the-floor'.

Contrastingly, the content of the high-frequency band usually consists of either steady-state sounds (e.g., synth or lead) or very short and rhythmic transients (e.g., hi-hat or percussion). The combination of a steady pulse and the low-frequency signals compressing high-frequency elements are typical for EDM.

A similar approach to the previous example was taken to mix this track. First the crossovers of the bands were adjusted following the values used in the previous experiment. Then the ratio and threshold were set to extreme values to strongly exercise compression. Finally, the input gain and the make-up gain were adjusted to balance the level of the frequency bands. The result was then bounced and normalised in MATLAB.

4.3.7.4 Commentary – Aural Analysis

As was predicted, pumping is more obvious in this mix compared to the previous architecture. However, it has a different characteristic from the sidechain and the nonlinear mix which makes it sound unusual. Normally, pumping a track either using a sidechain or a nonlinear mix will produce the behaviour of ducking the non-dominant elements. This typically results in the gain reduction of the high-frequency content of a mix. However, for the cross-band nonlinear mixing architecture, it sounded as if the level of the mid frequencies on the synth pad are boosted when the bass drum is present. The result gave the impression of an inverted sidechain.

It was discovered that the low-mids of Stem 2 that were sidechained to the high-mids of Stem 1 were dominant on the downbeat, which is when the bass drum is present. This is because the hi-hat only plays on the upbeat. Consequently, when Band 3 of Stem 1 is not occupied on the downbeats, the space is filled by Band 2 of Stem 2 which is the pad. Consequently, it unintentionally led to the impression of the bass drum boosting the levels of the pad when in fact it was the result of an absent element that made the pad louder.

Comparing this mix to the previous mix demonstrates the importance of the mid and high-frequency content in indicating amplitude modulation in a mix. For example, in the previous experiment, even though pumping did occur, it was not perceived to be as dramatic because the high-frequency content was at constant level most of the time. In this experiment, pumping can be perceived because Band 4 is subjugated to Band 1 and the ‘boosted’ Band 2 in Stem 2 can be clearly perceived.

There is a biological explanation for the common musical practice of using low-frequency content to provide temporal cues and high-frequency content to contain detailed spectral information relating to the pitch of both chords and melody lines (Hove, et al. 2014). Using electroencephalography (EEG), it was demonstrated that the humans’ auditory cortex detects pitch more robustly in the higher tones and has better temporal detection for lower pitched tones. Therefore, the effect of compression of the spectral content of a musical signal may be better detected by listeners if the magnitude of the high-frequencies are reduced by the dynamics processor, as demonstrated in this experiment. This means that triggered amplitude modulation may be more easily heard when it is in the high-frequency range.

The spectrograms of both mixes are shown in Fig. 4.6. The magnitude of the ambient pad (red arrow) gets higher when the bass drum (purple arrow) is present in the track. It gives the impression that the magnitude was boosted. However, that is not the case because the occurrence of hi-hat (orange arrow) reduced the magnitude of the pad as indicated by the reduced magnitude. When the hi-hat is absent, the ambient pad returns to normal. This is not the case for the result from the previous architecture as the spectrogram of the output shows no dramatic fluctuations.

It can therefore be concluded that for a multi-band nonlinear system, an alternate architecture may produce a better result if the goal is to produce a dynamic mix. An inter-band system does not impart noticeable amplitude fluctuations on the frequency bands owing to the nature of the signals in the frequency bands. This means that low-frequency content modulating one another will not produce noticeable pumping, whereas for high-frequency content, either the rhythms will be syncopated and varied, or the signals will be steady-state and therefore will not impart strong amplitude fluctuations.

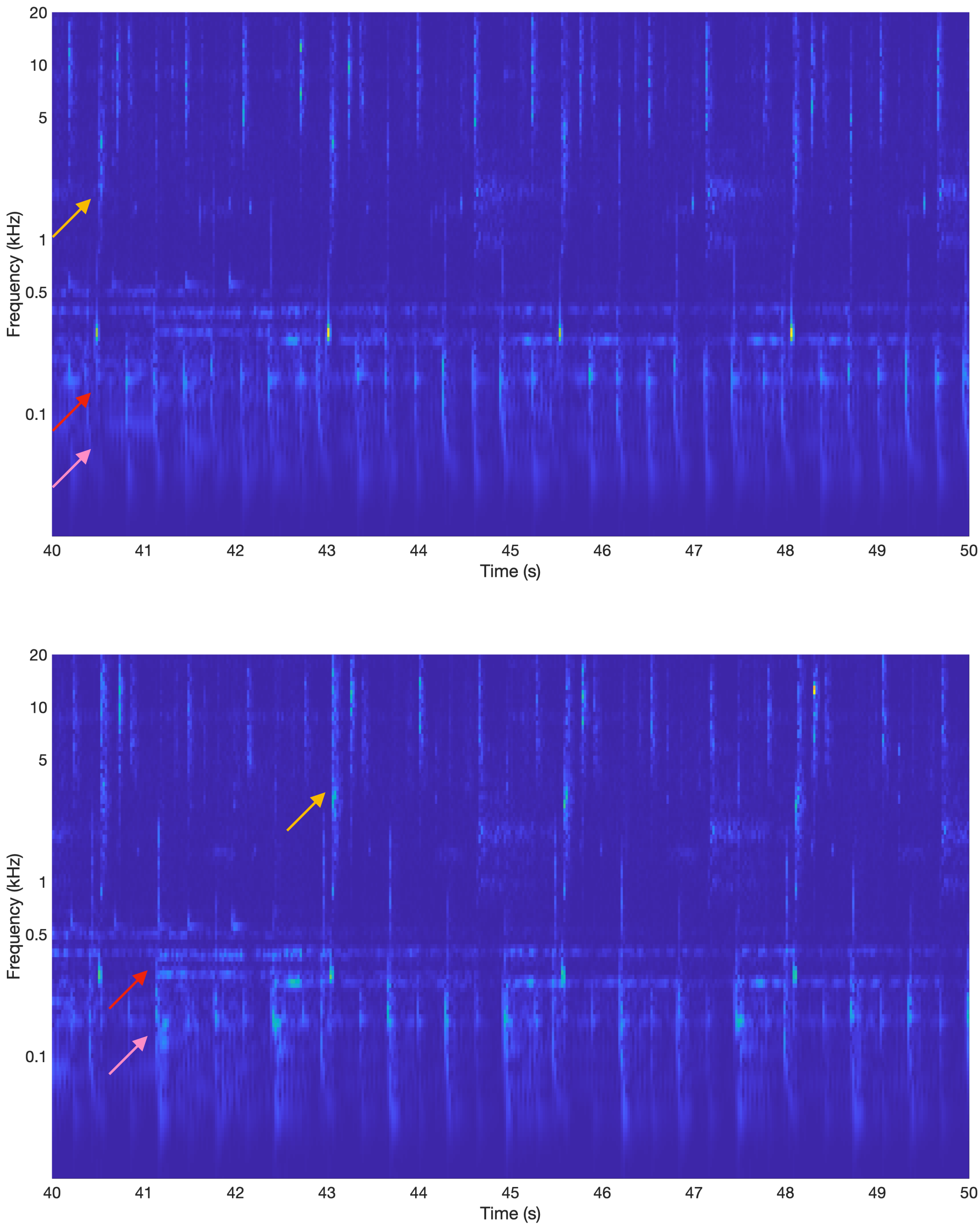


Fig. 4.6 : *Spectrogram of the inter-band nonlinear system (top) and the cross-band nonlinear system (bottom). Red arrow is the ambient pad, purple arrow represents the bass drum and orange arrow represent the hi-hat. Notice that the area pointed by the red arrow is brighter at the bottom image compared to the top.*

4.4 Discussion – Sidechain, Nonlinear and Multi-Band Mix

Comparing the result of the four architectures explored in this chapter, a few differences can be highlighted. In particular, the nonlinear mix and the sidechain mix demonstrate the following differences:

1. **An element will not be perpetually dominant in the nonlinear mix.**

   For the sidechain mix, typically one element will be assigned to be the dominant. Whenever the dominant element exceeds its threshold, it will trigger amplitude modulation of the other tracks that are in the system. This is not the case for the nonlinear mix because any elements that exceed its threshold will trigger compression for the other tracks in the system. Therefore, dominance is not assigned to one element and can be changed in an instant.

2. **The top-down mixing approach is more suitable for the nonlinear mix.**

   For the sidechain mix, either the bottom-up or top-down approach may be used. This is because changing the parameters in any of the subjugated tracks will not affect the others. However, for the nonlinear mix the interaction between all the tracks in the system must be taken into consideration while mixing the tracks. The reason for this is that the system is interconnected whereby any changes in one track will affect all the others.

The differences between the multi-band nonlinear mix and the nonlinear mix are not comparable to those between the sidechain mix and the nonlinear mix. This is because the difference in their architecture have produced a completely different result. Among the differences that can be observed are:

3. **The tone of the elements is affected.**

   For the nonlinear mix, only the amplitude of each track in the system is modulated. Thus, the tonal character of the elements is not affected. For the multi-band system, however, because the elements were split into four different bands prior to compression the spectral balance of each element was affected by compression. Thus, tonal features of the elements will change.

5 New Use Cases

5.1 Introduction

Thus far, this research has proven the hypothesis on the characteristics and the behaviour of a nonlinear mixing system created in Max 8. From the tests conducted using music excerpts, it was shown that the nonlinear mixing system can be used both as a mixing tool or applied creatively in a composition. One example is that clarity in the mix can be improved by reducing masking, using subtle compression. Additionally, applying overt compression will produce interesting textures, resulting from the interaction within the tracks in the system. These tests, however, have been conducted by only one user who is the creator of the system.

Considering that the architecture provides the opportunity for it to be used to achieve several different results, at this stage it is important to investigate how other users would apply the system in their production. There are two main scales of use of this system; the first one is between subtle and overt, and the second is between corrective and creative. It is likely that these two are correlated in some way, whereby subtle is considered corrective and overt is considered creative. However, this is not necessarily always the case as a user may use subtle parameters for creative effect and vice versa.

It is also beneficial to gather opinions on the usefulness of the nonlinear mixing system. Furthermore, there may be other creative ways of applying this system in production that the author of this thesis might not have discovered yet. Thus, to discover the possibilities that may emerge, a small-scale experiment that has involved multiple participants has been conducted.

5.2 Deduction

Based on the author’s experience, the process of setting the parameter values is intricate because a small change on one element can affect the overall tonal balance. This may discourage participants from using extreme values. Applying subtle compression may produce results that sound more congruous with most of the typical current popular production techniques. Furthermore, there may be many producers who do not seek to create the level of novelty in their productions which is possible in this tool (e.g., Flying Lotus whose work was cited earlier in this thesis).

However, this does not mean that it will be impossible to observe new discoveries. This is because there may be other ways to apply this algorithm creatively in a production that have yet to be explored. Therefore, the author believes that it is likely that most of the users might use the nonlinear mixing system as a mixing tool, or more specifically, for corrective purposes.

5.3 Research Design

The data that needed to be collected were: (1) how the users typically use side-chain compression, (2) whether the participants used the tool to achieve creative results or for corrective purposes, (3) whether they have discovered novel ways of implementing the system in their mixing process that are not possible in more conventional DAW, and (4) the nature of any advantages of using this system for mixing. Accordingly, an interview which consisted of a mixture of close-ended and open-ended questions was planned. For this reason, a mixed-method approach using a combination of structured and semi-structured interviews was taken.

Structured interviews involve all the respondents being asked the same questions in the same order and they were required to select one answer from a range of options, hence the name structured. Semi-structured interviews are slightly similar to structured interviews whereby all the participants will be asked the same set of questions; the difference is that they are allowed to respond to it as they choose. The data that have been collected from the structured interviews are quantitative while the semi-structured interviews are qualitative (Morse 2012).

In terms of the research design, this experiment has used the explanatory sequential mixed method (Creswell and Creswell 2018). This means that the researcher initially conducts quantitative research, evaluates the results, and then uses qualitative research to expand on the findings and explain them in greater depth. It is deemed explanatory since the qualitative data help to explain the initial quantitative data results, and because the quantitative phase is followed by the qualitative phase, it is termed sequential. For this reason, in this interview the structured questions were followed by a semi-structured question.

The sample size can range from 2 to 25 participants or until theoretical saturation occurs (Beitin 2012). Theoretical saturation is the circumstance when the researcher has conducted enough interviews to feel as if he or she has learnt everything possible from the interviews and has double-checked those understandings by reinterviewing the most reliable and knowledgeable informants (Johnson and Rowlands 2012). However, this is also dependent on the resources (time and funding) that are available to the researcher (Beitin 2012).

The available resources for this experiment, particularly during the pandemic, have enabled fifteen people to be interviewed. There was consistency between the answers that were given, suggesting that a representative sample has been found. Furthermore, the interview was mainly structured, and the number of questions was small. Therefore, the sampled data is considered representative. The candidates for the experiments were screened prior to the experiment so that only those with the relevant criteria to the research will be selected. This has been done using a screening questionnaire.

The first criterion was that the participants must have at least one year of experience in music production. This is because the user needs to have a basic understanding of signal flow, compression, and side-chained compression. Thus, possessing at least one year of experience in mixing as well as music production should be enough to build an understanding on the details that have been mentioned earlier.

The second criterion was that they must use an Apple computer. This was because the system has been built and compiled using MacOS and is currently not supported by the Windows operating system. Apple computers are typically used in music production environments and so this is not deemed to be a particular problem.

The third criterion was that the participants must have access to a high-speed internet connection. Due to the movement restrictions that were enforced worldwide during the pandemic, it was impossible for this experiment to be conducted in person or at a predetermined location. Therefore, the only option was for it to be done online via Google Meet video conference. In order for the interviews to be held uninterrupted, the participant's broadband speed must be at least 3.2 Mbps, which is required by Google Meet.

The fourth requirement was that the respondents had access to a private working space where there are no distractions such as pets, television, or children. This is to ensure that the participants will be able to focus on completing the task that is given, to explore the system to its full potential and are able to understand how it works and how it is different from the conventional linear system.

After the participants were selected, they were provided with a link to download a file that comprised the nonlinear mixing system, eight audio samples, a user guide and an instruction that explained the task that they needed to complete, which was to use the nonlinear mixing system. No additional information was given regarding the approach in mixing, such as whether to use extreme or subtle parameter settings. They were also given the freedom to use their own samples (this however was not explicitly mentioned to the participants). This was to encourage the participants to develop their own approach in using the system, thus, providing unbiased feedback.

The participants were given ample time to complete the tasks by allowing them the freedom to choose when to conduct the interview. The sessions were conducted one person at a time rather than in groups. This has produced answers that are without external influence through discussion or overhearing the answers that are given by other participants.

5.4 Questions

There were eight questions in the questionnaire which can be categorised into four sections. The first section aimed to investigate the participants' experience with side-chain compression. This might give some indication regarding their understanding of side-chain compression, how they usually utilise it and finally highlight the differences between their typical usage of side-chain compression to the approach on using this architecture.

Then, the interview focused on how they applied the nonlinear mixing system, whether subtle or overt. These questions were closed-ended, whereby the participants were only required to give answers based on the list that was given by the researcher (e.g., agree or disagree; corrective or creative). This was then followed by the open-ended questions, to explore whether they had opinions on the advantages of using this system or if they used the system in a way that was different from how they mixed in their DAW. Data analysis was then conducted at the end of the collection.

Question		Type	Purpose
1	Do you agree with this statement? I know how to set up side-chain compression in my DAW.	Structured	To find out how the participants typically used side-chain compression.
2	Do you agree with this statement? I have applied side-chain compression in my previous project(s).	Structured	
3	In your previous project(s), did you often use side-chain compression for corrective purposes or creative purposes?	Structured	
4	For this experiment, did you use the nonlinear mixing system to create pumping effect or to reduce masking?	Structured	To find out how the participants used the nonlinear mixing system.
5	Do you agree with this statement? There is an advantage/are advantages of using this system.	Structured	To gather opinion.
6	What is the advantage/are the advantages?	Semi-structured	
7	Do you agree with this statement? I discovered a mixing method/technique that is unique to this system.	Structured	To find out novel application in mixing.
8	What is the method/technique?	Semi-structured	

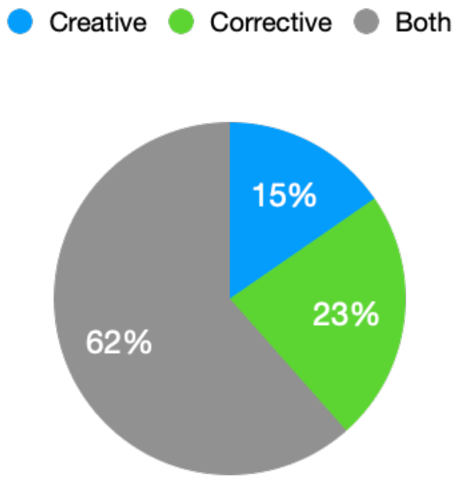
Table 5.1 : Questions for the Interview.

5.5 Interview

The online interviews were conducted over the course of four weeks and involved 15 individuals. All of them were music producers with at least one years' experience in music production- either in a recording studio, post-production, broadcasting, or in academia. They were recruited online through various social media platforms. Most of the respondents were Malaysians, except for a few individuals. Some of the respondents were not proficient in English; in this situation, the answers were given in Malay and translated post-interview.

5.6 Result

5.6.1 Part 1 and Part 2



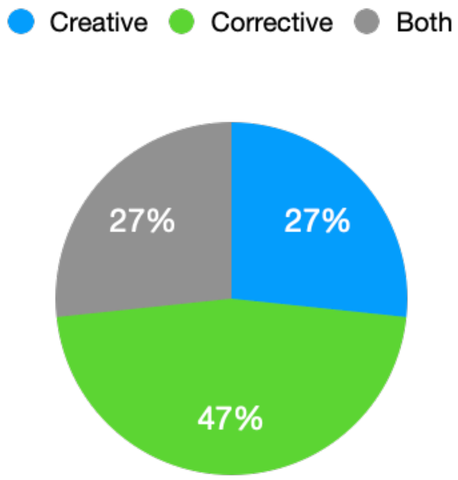


Fig. 5.1 : *How the participants typically used side-chain compression (left) and, how they used the nonlinear mixing system (right).*

All the participants knew how to set up the sidechain compression in a DAW, although some have never used it in their production before. From the interview, we know that most of the respondents who have experienced using this routing system in their previous projects usually applied it for both creative and corrective purposes. However, there are also a significant number of them who have only used it for creative effect.

Surprisingly, for this experiment a majority of the participants only used the nonlinear mixing system to reduce masking. Most of them commented that it is easy to achieve clarity using the system. Nevertheless, one subject pointed out that it is slightly difficult to get the right balance regarding how much compression should be applied. This is because any changes on one track will affect all the others, which may be frustrating when a small variation on one track affects the balance that the user strived to achieve.

5.6.2 Part 3

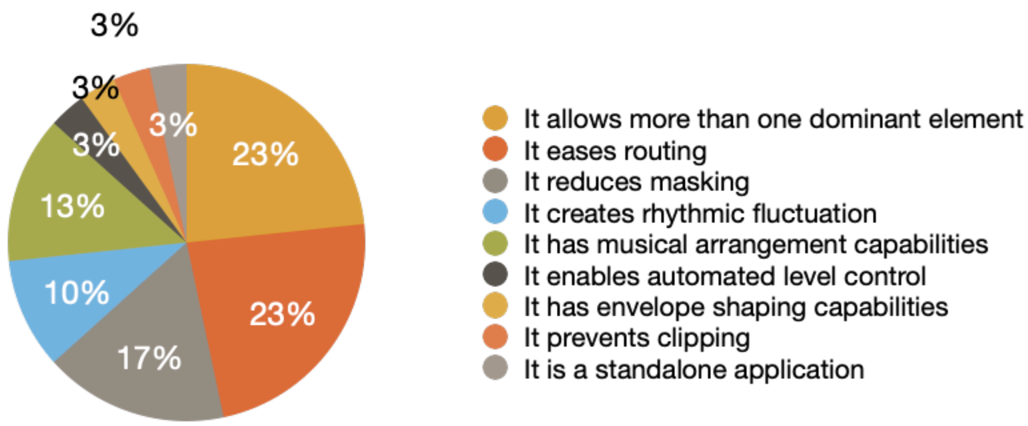


Fig. 5.2 : *The advantages of using the nonlinear mixing system.*

All the individuals in this experiment agreed that there are advantages of implementing the nonlinear mixing system in their production. Over twenty percent of the test subjects mentioned that using buttons to engage or disengage sidechaining made it easy for them to implement it in their mix. This means that they do not have to manually route the tracks because everything is already inherently connected to one another. This feature accelerated their decision-making process because they could then just simply choose which element is to be dominant. One user mentioned that less time was spent thinking about the routing, offering more time to focus solely on getting the right sound. According to the testers, allowing for more than one element to be dominant is also beneficial to their workflow since this broadens up more possibilities in mixing as it offers flexibility. The users are not restricted to having just one track controlling the dynamics.

In line with what was reported in this thesis, a significant number of the subjects commented that the system reduces masking and gave the elements in their mix more clarity. One user noted that the system's unique ability to shape the temporal envelope of sounds was quite useful to make the samples to fit one another. This produces clarity in the mix whereby the tail of an audio sample can be slightly attenuated to allow the attack from other tracks to be more audible.

The use of the nonlinear mixing system as an arrangement tool was also noted. A few of the participants commented that the volatility of the system may produce unexpected results that might be useful and may inspire creativity.

Other than that, the participants also commented on the advantage of utilising the system to produce rhythmic undulations. However, this comes as no surprise as it is currently a popular use of sidechain compression. One of the participants noted the advantage of having the system as a standalone application without needing to load it in a DAW; this made it effortless for him to implement this system in his project. Finally, only one user noted the advantage of using the architecture to assist in level control and as a means to prevent clipping.

5.5.3 Part 4

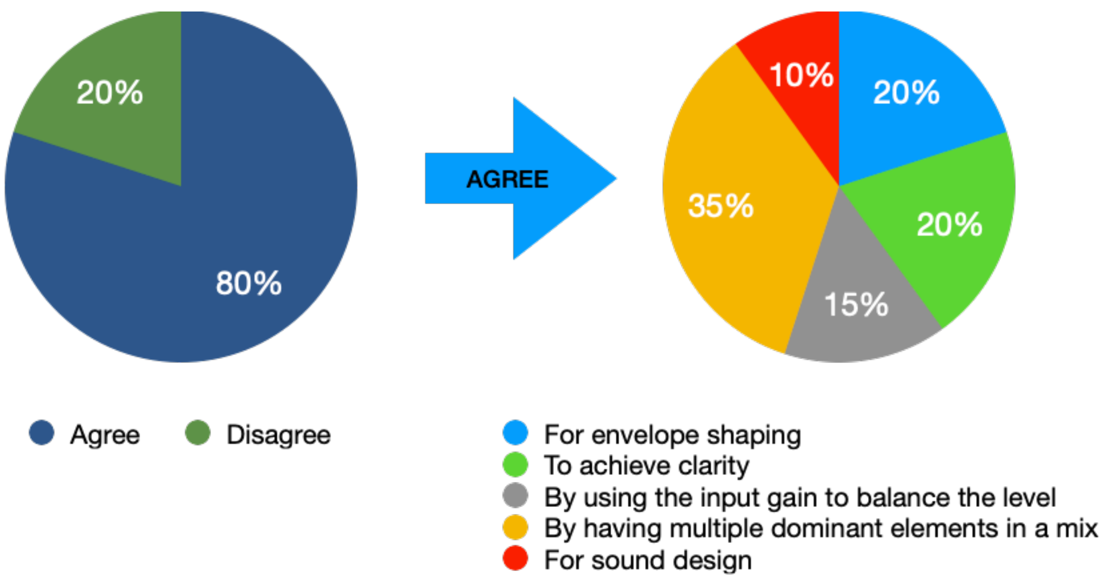


Fig. 5.3 : *Percentage of participants who discovered novel approach of mixing using the nonlinear mixing system (left) and their discovery (right).*

Eighty percent of the participants agreed that they found at least one mixing technique that is unique to this system. From this amount, most of the respondents took advantage of the system's inherent ability to allow more than one dominant element in the mix. This ability, for them, was a new discovery that opened more creative possibilities.

As a result, the users found unique ways to apply the system by utilising its ability to provide clarity in their mix. One of these is by using multitimbral bass in a pop or EDM song. For example, one user commented that the clarity that this system produces encouraged them to use more than one low-frequency element in their song. Similarly, another producer who specialises in trap and hip hop noted that all the low-frequency elements in the nonlinear mixing system can be clearly heard even though they used quite a lot of bass samples in their experiment.

Furthermore, the participants noted its unique envelope shaping capacity. One user utilised it in a drum bus, using the kick, snare, and toms to trigger compression on the overheads for clarity. Even though this can also be achieved in a DAW using bus compression, the routing that needs to be done in order to achieve the same result might be complicated for non-experts.

Another unique way in which the nonlinear mixing system was implemented was by using its input gain to instantly change the balance of the mix. By just making changes to this parameter, the level of the other tracks can be automatically attenuated. The user commented that this is different from their normal mixing practice because they do not have to adjust the volume of all the individual tracks to obtain similar results. Another user likened it to a 'lazy-mode sidechain compression system'; this participant made a remark on how the system can be used to automatically attenuate any elements that are not the focus in the mix by adjusting the input gain.

The system was used as a sound design tool, whereby a participant uploaded different notes that make up an arpeggio on the different tracks and used the compression to shape the envelope of each note. Another user applied it to create an interesting rhythm on an ambient synth pad and exported the synth pad to be used in their production.

5.7 Discussion

The result from the interview indicated that nowadays, audio engineers are more inclined to use side-chain compression as a creative tool in their production. This is highlighted in the first section of the interview whereby only 23% of the participants used it solely for corrective purposes. However, this is not the case for the nonlinear mixing system as almost half of the percentage of the respondents utilised it only to reduce masking.

As was mentioned by one of the interviewees, it can be frustrating to have the balance of the mix change drastically when an element is slightly modified, which is the case for extreme compression setup in the nonlinear mixing system. Furthermore, the undulation caused by compression may sound random as it is triggered not only by a steady pulse (usually the kick bass in the case of electronic dance music), but a mixture of all the elements in the system. These behaviours may discourage audio engineers applying extreme parameters on the processor. Therefore, subtle compression may be the preferred implementation for most users.

On the subject of the systems' advantages, the responses were split between the corrective and creative use of the nonlinear mixing system. Even though in the previous section almost half of the participants only used it to reduce masking, they did mention other ways of utilising the system creatively. It is therefore the authors' opinion that this can be studied further by using different samples or compositions that specifically explore the creative potential of the nonlinear mixing architecture.

This is reflected by the results in the next section whereby one fifth of the respondents did not find (or have not yet found) a novel way of mixing by using the nonlinear mixing system. This might be because their creativity was limited to the samples provided for the experiment. As users integrate this tool into their own workflows outside of this experiment, it is likely that they will feel more able to select materials and work with them in a more iterative way.

5.8 Conclusion

This experiment has investigated different use cases of the nonlinear mixing system amongst a variety of users. Despite some limitations while conducting this experiment due to restrictions caused by the pandemic, the author believes that this study offers answers to key questions. The result from this study showed that users will most likely utilise the system as a mixing tool to reduce masking. Secondly, most of the participants agree that there are advantages of utilising the system in their mix.

Creative uses of new technology do not happen overnight. Therefore, as more users utilise this tool in their production, other novel ways of benefiting from this tool may be discovered. It has taken more than half a century from the year when compressors were first invented to the point when their creative use in music production has become mainstream. Thus, it may be the same for this architecture whereby other creative ways in which this nonlinear mixing system would be used in production may emerge.

With the signal processing properties of the system established, personal use cases explained by the author and others investigated with fellow producers, the next chapter will recapitulate, discuss, and finally summarise the findings of this research.

6 Conclusion

6.1 Summary

This research has developed a novel architecture which has then been expanded to produce three different configurations. To provide a comprehensive understanding of how the system was developed and a clearer understanding of its purpose, this research has focused on three main areas. The **first chapter** provided an introduction for this research followed by the **second chapter** that proposed using the CQT spectrogram to analyse the spectral content of a mix. This thesis argued that the STFT is not necessarily the optimal method for presenting time frequency information about these signals. This is because the CQT provides better resolution for low-frequency content and better temporal resolution for high-frequencies.

A key factor in existing sidechain compression systems relates to the configurability of the audio processing environment. The digital audio workstation routing capabilities and limitations were investigated to understand how a massively mutual sidechain compression system might be implemented. This was because it was demonstrated that the routing for such systems using analogue workstations is highly complex because of the cabling and the number of busses needed. Such routing in a DAW does not require any cables. However, feedback is not possible. Ultimately, neither of these domains were able to produce the behaviour of the system described and investigated subsequently in this thesis.

This work then continued by exploring the history of compression, from the analogue processors using valves, photo-sensitive resistors and finally transistors to its digital adaptation using computer software. Here, it was clear that the use of dynamic range processors changed from corrective to creative. Initially, it was used to prevent malfunctions on broadcasting stations and nowadays, as was described previously, compressors are also used for aesthetic reasons. Various available architectures of digital compressors were afterwards explored which then led to proposing a novel way that it can be used to build a system.

After the theoretical groundwork had been covered, the **third chapter** contributed to the developmental section of this thesis. First, a bilateral compression system was theorised by producing a block diagram, a processing flowchart and a mathematical representation of the system. It was then expanded to include more than two tracks to create a nonlinear mixing system. Neither architecture had been implemented prior to this research. By understanding the algorithms involved in the process, a few deductions were made that served as a guideline to verify a successful implementation of the bilateral processing architecture. The identifiers are: three distinct sections indicated by the presence of two corresponding knees in both test signals; the presence of sidebands caused by cross-modulation; and, the correspondence of the output of the system to the type of signals used in the test (whether they are transient or steady-state sounds).

This two-track bilateral compression architecture was then implemented in two chosen DAWs; LPX and Reaper. However, the results did not produce the predicted outcome. This is because mutual feedback architecture is not supported in either, indicated by the lack of sidebands in the frequency magnitude plot from the output. Therefore, this system was implemented in the visual programming language Max 8 using gen~ in order to ensure that feedback loops will only have the latency of one sample and the recursion is near-instantaneous.

First, a two-track bilateral compression architecture was created. It was then examined using a few types of test signal to verify the correct implementation of the algorithm. The results of the tests were positive and served as templates for further tests which consisted of a real-time application of the system and exploration of the behaviours under different parameter settings. Following this, two more variants of the two-track bilateral compression systems were created. The first was an extended version of the architecture which is an eight-track nonlinear mixing system and the second one is a two-track multi-band nonlinear mixing system. The latter utilised Linkwitz-Riley crossovers to split an input audio signal into four bands before processing each frequency band using bilateral sidechain compression.

The **fourth chapter** provided some musical examples to demonstrate the applications, capabilities and potential for new musical expressions of the nonlinear mixing system. The first example was a piece of ambient music which consisted of five transient audio signals and three sustained (or steady state) sounds. Three variations were provided using different mixing approaches; a nonlinear mix (audio example 1), a linear mix (audio example 2) and another mix created using sidechain compression with the bass drum as the dominant element (audio example 3). For the nonlinear mix, the intention in the production was to induce overt pumping. The linear and sidechain mix were then produced to closely resemble the output from the nonlinear system. All the results were then normalised, analysed and compared.

Not only do the results show differences in the character of the amplitude modulation, but aural analysis also revealed that the nonlinear mix sounded more spacious compared to the other two mixes. Timbral variations was also discovered to be overt in this mix and may be used to attenuate masking. To substantiate this claim, the time frequency representation derived from the CQT for this mix was analysed which exhibited slight gain reduction on the overlapping frequencies.

Another representation of compression behaviour was presented, showing how the level control signal varies over time. This approach was taken because the amplitude modulation was not immediately and clearly discernible by the ear. The control signal plot helped to identify the detailed interactions which create an overall audible impression. The resulting plot shows complex gain reduction patterns. These complex behaviours, although seemingly erratic, can impart a useful variation onto a sound.

For the next audio example, an experimental hip hop piece was mixed utilising the nonlinear system. The track consisted of 6 transient signals and one sustained pad. Three mixes were also produced; the first one was conducted using the system developed in this research set to overt compression (audio example 5), the second one is a linear mix (audio example 6) and finally a sidechain mix (audio example 7). For the nonlinear mix, the aim was to introduce overt compression therefore the threshold and ratio were first set to the extreme values. Then, the linear mix and sidechain mix were produced.

Aural comparisons of the nonlinear and the sidechain mixes highlighted a few characteristics. The amplitude fluctuations of the bass drum in the nonlinear mix gave the impression of timbral variations due to a masking effect. Another behaviour that was highlighted was that the top-down mixing approach was found to be more suited for this processor. This characteristic provides the users the feeling of live performance in mixing because they will have to listen to the whole tracks while mixing, unlike a typical mixing session whereby they may listen to one track at a time while balancing the level. The impression of live performance and human manipulation of parameters is enhanced by the complex gain reduction pattern, unlike uniformed gain reduction determined by the dominant bass drum for the sidechain mix.

Discussed in the final section of this chapter were two more mixes produced using the multi-band non-linear mixing system. Using the same samples as the previous experiment (experimental hip hop), the audio tracks were first grouped into two stems and summed after which it was loaded onto the system. The first stem consisted of drum parts and the second one is the other tracks.

For the first multi-band nonlinear mix (audio example 8), the system only connected the same frequency bands to one another through bilateral compression then the parameters were set to the values that produced the result which is preferred by the author. Because bilateral compression only occurs between the same frequency bands, pumping which is usually by the low-frequency content now only happens in the low-frequencies (Band 1). The high-frequency content of the ambient pad also became more apparent. It was found that the content of Stem 1 (hi-hats and shakers) was not powerful enough to cause modulation on Band 4 of the ambient pad (Stem 2).

It was also discovered that the mix sounded lifeless and dull even though pumping can be observed in the low-frequencies. The proposed explanation was that the high-frequency content of the output was passive and not dynamic. Thus, it is proposed that the dynamic variation of the high-frequencies are important indicators for a song to sound energetic and exciting.

The final audio example (audio example 9) was produced using a variation of the multi-band nonlinear mixing system. For this architecture, rather than connecting the same frequency bands to one another, it was set to have cross-band processing capabilities. This produced an output that sounded entirely different from the previous mixes. For a start pumping is more obvious in this multi-band mix compared to the previous one. This behaviour supported the previous proposal regarding the importance of the mid and high-frequencies in signalling amplitude modulation in a mix.

One character that made the mix sound different from the eight previous versions is that the ambient pad sounded as if it was boosted by the bass drum instead of attenuated. The cross-band processing has led to the third frequency band of the first stem modulating the second band of the second stem. Because the hi-hat was programmed to only play on the upbeat, it gave the ambient pad space to dominate the downbeat, thus giving the impression of the bass drum playing on the downbeat affecting the synth.

The final section of the third chapter was a discussion that summarised the behaviour observed in the test. The first two observations were about the nonlinear mixing system. The first was that there will not be any one element that will be perpetually dominant in the mix unlike in the sidechain system. The second observation was that a top-down mixing approach is most likely to be the best approach when mixing using the nonlinear system. All of these behaviours are unique to these processing algorithms and, to the best of the author’s knowledge, have never been observed in other mixing systems prior to this research.

Finally, in the **fifth chapter** of this thesis the nonlinear mixing system was put to use by 15 users in a small-scale experiment. The participants were provided with the same music excerpts used by the author in the previous experiment and were instructed to mix the samples using the tool developed in this research. Due to the restrictions caused by the pandemic, interviews were conducted online via Google Meet. A mixture of structured and semi-structured questions were asked to investigate how the participants would likely utilise the tool in their mixing environment and the nature of the advantages (if any) of using this system.

The study found that users mostly utilise it as a mixing tool, or specifically, for corrective purposes. All the participants agreed that there are advantages of using the nonlinear mixing system. Furthermore, a majority of the interviewees have discovered mixing methods that are unique to this system.

6.2 General Discussion

In this research, I have successfully implemented a two-track bilateral compression system which was then expanded to form an eight-track nonlinear mixing system, a two-track multi-band nonlinear system and a two-track cross-band multi-band nonlinear system. The first architecture enables eight stereo tracks to simultaneously attenuate one another via sidechain compression. The second and third system split two stereo tracks into four frequency bands which were then used to mutually compress one another.

The cross-modulation algorithm in this system allows for bilateral processing whereby a signal can alter not only the dynamics of another but can also be attenuated in return by the same signal that it controlled. The recursive nature of this system is unique to this research, partly because DAWs do not have the capability for this kind of algorithm. The other reason is that routing takes too much effort to create this using analogue processors. Therefore, the architecture was implemented successfully by programming it from scratch using Max 8.

Tests presented in this thesis demonstrate that this system functions as expected. The architecture was then expanded to be able to process eight tracks simultaneously and also by implementing a band splitting algorithm to turn it to a multi-band processor. These systems not only opened the door to experimentations, they also inadvertently highlighted a few details that may have been overlooked.

One of these details is the importance of rhythmic undulations occurring in the high-frequency range of a song. Without any fluctuations it may not be more interesting than static noise on television. Another detail is the mixing approach: for a massively interconnected system where a change in one will affect all, the top-down mixing approach is most likely to be appropriate.

The author of this thesis has personally utilised the system in his projects (non-research related), particularly in situations where the elements are within the same frequency range or have similar timbre. One example was to mix a recorded *sapé*[11] trio whereby the performers were playing simultaneously within the same octave range albeit with a slightly different rhythmic sequence. This arrangement was problematic because the three recorded tracks were masking one another. One of the ways audio engineers typically solve this issue is by positioning the elements in a different stereo field using a standard panpot. However, using extreme values will make the mix sound disjointed as if the musicians are not performing in the same space. Conversely, subtle panning can cause masking.

This was one of the situations where the nonlinear mixing system proved to be beneficial. The compression provided by the tool improved the rigidity of the attack of every *sapé* in the song. The sustain of the strings were also slightly ducked by the transient which provided clarity for each element in the arrangement. These characteristics allows panning to be done moderately. This result cannot be achieved using a typical side-chain compression architecture because it does not accommodate multiple dominants in the system.

[11] The sapé is a traditional music instrument from Borneo. It is a four-stringed chordophone and played by plucking the strings using fingers.

Having said that, this is only from the author's personal experience and following contemporary production techniques. It may well be that this novel architecture might open new paths of creation that are not currently possible. This may be a useful tool for music producers that are looking to explore new sounds and textures and to assist in achieving clarity in a mix. Moreover, producers in the future may use it in ways that I have not anticipated, as has happened with other signal processing tools throughout the history of music production.

6.3 Future Work

This research serves as a proof of concept for a successful implementation of a massively mutual compression system. It has also demonstrated how such a system can be applied in production, correctively or creatively. However, it might be limiting to use such a processor in a standalone environment. Therefore, it would be in the best interest of producers to study the feasibility of this architecture to be fully integrated in commercially available DAWs.

This raises one fundamental question; how would this work? The first thing to be considered is that it would need DAW developers to build it into their system, which would require sample-based processing. This would also require computers to have computational power fast enough not only to process all the audio effects typically used in production such as equalisation, reverberation et cetera and to trigger MIDI samples, but also to process simultaneous compression that occurs in all the tracks. As the block size is reduced in the traditional CPU based DAW, then its processing efficiency declines. Therefore, either faster processing is needed, or a change to the architecture of the computing systems which are typically used to host DAWs might be needed.

Ultimately, this is almost making a case whereby the predominance of general-purpose computers like laptops for the purpose of making music is challenged because all of a sudden there are things that the computers are not able to do. We have been told that digital platforms enable us to create or do whatever we want but, in this case, it is very difficult. This system had to be created in gen~ as a standalone application instead of implementing it within a DAW. This required programming and signal processing knowledge. If the system were to be designed from scratch, then it would be possible to deploy digital signal processing hardware which would be better capable of dealing with signals on a sample-by-sample basis.

The issue of whether it is worth investing to integrate such architecture in their workstations is left to be answered by DAW developers and hardware manufacturers. On one hand, this may be the feature that differentiates them from competitors which can be a selling point. On the other hand, they might have to make modifications to their source code which may be too much work for too little return. But it remains the case that these architectures, as they currently exist, do considerably limit the routing freedom within them. They may even have to develop new processing hardware just to readily support such an architecture.

One outcome that cannot be ignored however is the impact and the possibilities offered by the nonlinear system in music production, especially in experimental genres and EDM. This is because, both these genres seem to be at the forefront of new sounds and musical textures in which a novel production approach definitely will be of their interest. The DAW that first implements a nonlinear mixing architecture developed in this research may be the game-changer that brings forward a new approach to making music.

This thesis also highlighted a few possible ways for this architecture to be implemented in music production. One of it is for the system to be used as a form of assisted mixing environment. Creating a database of parameters (or presets) for different uses such as specific music genres or different types of signals (e.g., transient or steady-state) may improve the nonlinear mixing system for this application. Other than that, it can also be used for sound design. Musical works with complex undulations can now be created using a tool which allows for a more nuanced control. This, as a result, may produce interesting textures which might never have been achieved before. For further experimentation, the author of this thesis will collaborate with other composers to create works specifically designed to fully explore the creative use of the nonlinear mixing system.

6.4 Final Thoughts

Throughout the course of this research not only have I understood more about the creative application of the dynamic range compressor, but I have also developed as a musician and broadened my understanding of the theory and practice of music production. The reason for this is that I have come across more creative possibilities to produce work that contains more musical elements without worrying too much about masking, particularly in the treatment of drums and percussion. Furthermore, knowing how to manipulate the low-frequency content in my production has made my music more aesthetically pleasing.

This is made even simpler using the system that was developed in this research. Using this system, sidechain compression can be activated just with a click of a button. Consequently, the consideration of whether to use this processor in production or not is made easier because it does not involve setting up the architecture needed in a DAW to achieve the effect. Not only does this save time, but it also allows inexperienced users to utilise creative compression in their composition. One disadvantage is that at the moment it is not integrated in a DAW.

This research may be used as a tool for more discoveries and innovations to be made in music. What might be considered a misuse of an audio effect now may well be widely accepted in the future. It is hoped that this thesis will encourage more research in the creative use of audio effects and promote the creation of new technological possibilities in music production.

Appendix 1

This appendix provides the user's manual for the nonlinear mixing system and the multi-band nonlinear mixing system. To run the system, the user must first install Max 8 on their computer.

The images are the user interface for both systems. On it are red coloured numbers which corresponds to the numbers in the table on the page following the image. For example, to load a track on the system the user will have to click the button labelled "Load Track".

To record the output from the system, the user must first click the "Rec" button (button 11) followed by "Start" (button 12) before clicking "Play" (button 9). The audio will automatically be saved to the same folder that stores the audio source loaded onto the system.

The standalone application for both systems are also available online.[12]

[12] https://kajibunyi.wixsite.com/hairulhasnan

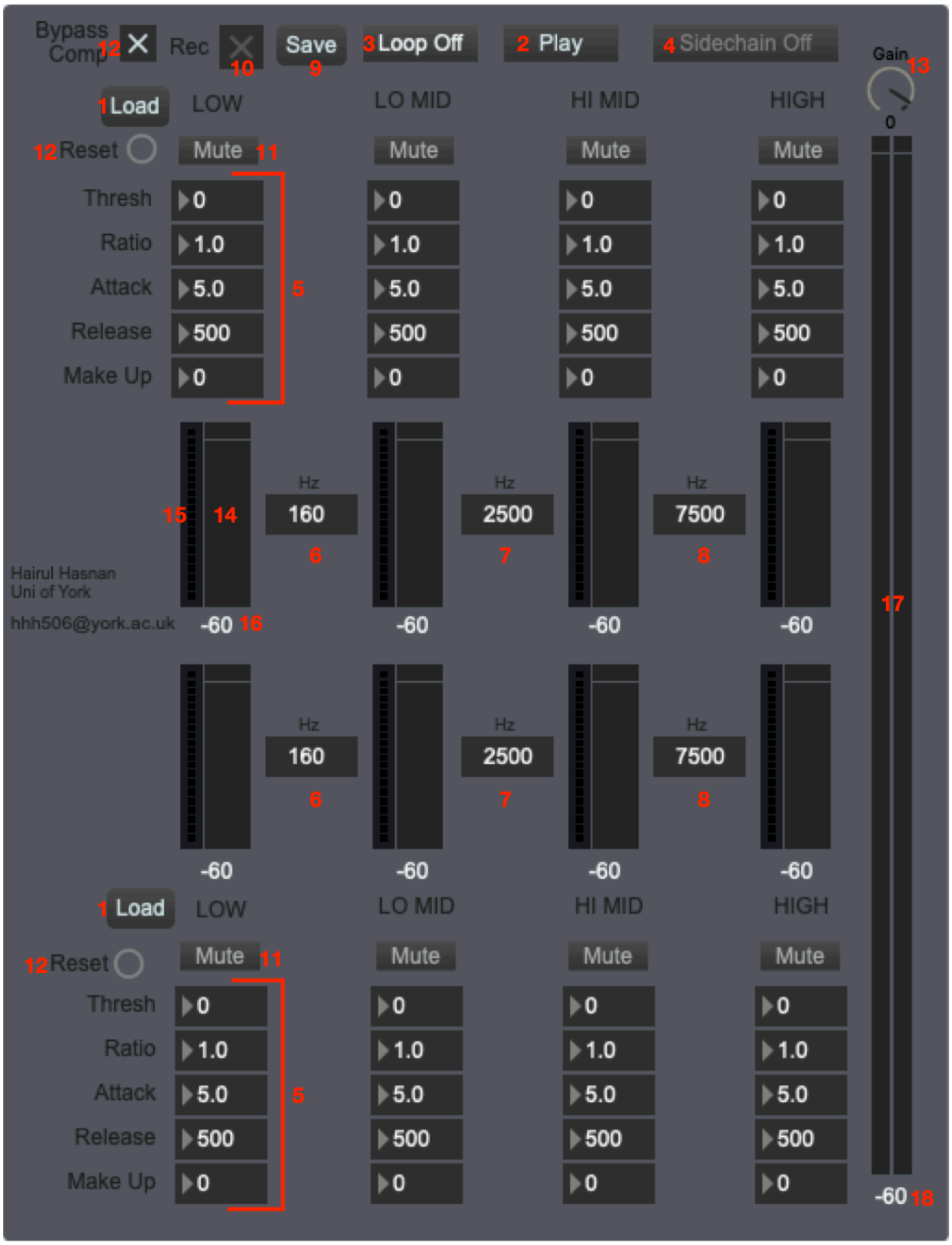
Nonlinear Mixing System

User Guide

Number	Label	Function
1	Load Track	Click button to load audio signal to the track. Supported formats : wav, aiff and mp3.
2	Loop On/Off	Click button to loop audio signal. The text on the button will turn white if loop is triggered.
3	Play/Stop	Click button to play or stop audio signal. The text on the button indicates whether button is set to play or stop the audio track.
4	Label is on the left	Click and drag mouse up or down to increase or decrease the value of the parameters.
5	None	Peak meter indicating the peak level of the track.
6	None	Numerical indicator for the maximum peak value of the track. Click number to reset value.
7	SC	Sidechain on/off. Button turns green when the track is routed to the nonlinear mixing system, or dark grey when it is routed to a normal compression system.
8	M	Button turns blue when the track is muted. The track still contributes to the nonlinear mixing system even when it is muted.
9	Play/Stop	Click button to simultaneously play or stop all the tracks.
10	Save CV	Click button to record the control signal of the compressor for all tracks for analytical purposes.
11	REC	Click button to record the output of each track, including the summed output.
12	Start/Stop	Click button to start or stop recording the output of each track, including the summed output.
13	Label is at the bottom of the button.	Click to reset the parameters for all tracks to default (default values are shown in the image).
14	Bypass On/Off	Click to bypass the system, only sending the original tracks without any processing to the master out.
15	Master Gain	Click and drag mouse up or down to increase or decrease the value of the output gain on the master out.
16	None	Peak meter indicating the peak level of the master out.
17	None	Numerical indicator for the maximum peak value of the master out. Click number to reset value.

Multi-Band Nonlinear Mixing

User Guide



Number	Label	Function
1	Load	Click button to load audio to Track 1 or Track 2. Supported formats : wav, aiff and mp3.
2	Play/Stop	Click button to start or stop both tracks simultaneously.
3	Loop On/Off	Click button to play audio tracks in loops. Text turns white when loop is activated.
4	Sidechain On/Off	Click button to send tracks to bilateral sidechain system or normal compression. Text turns white when side chain is activated.
5	Label is on the left.	Click and drag on the numbers to change the values of the parameters for compression.
6	None	Click and drag on the numbers to change the crossover frequency for Band 1.
7	None	Click and drag on the numbers to change the crossover frequency for Band 2.
8	None	Click and drag on the numbers to change the crossover frequency for Band 3.
9	Save	Click button to create a file to save the output from Track 1, Track 2 and master out.
10	Rec	Click the ‘X’ button to start recording the output from Track 1, Track 2 and master out. Click again to stop.
11	Mute	Click button to mute the frequency band.
12	Bypass Comp	Click ‘X’ to bypass the system and only sends the original input to master out.
13	Gain	Click and drag up/down to increase/decrease the gain for the master output.
14	None	Peak amplitude meter for each band.
15	None	Compression indicator for each band.
16	None	Numerical representation for the peak dB level of the band. Click number to reset value.
17	None	Peak amplitude meter for master output.
18	None	Numerical representation for the peak dB level of the master out. Click number to reset value.

Appendix 2

Participant Screening Questions

1. How long have you been involved in mixing or music production.

[ ] Less than 1 year	[**Terminate**]
[ ] 1 - 2 years	[Continue]
[ ] 2 - 5 years	[Continue]
[ ] Over 5 years	[Continue]

[**Recruiter**: Recruit a mix.]

2. Do you use a personal computer with Windows or Mac operating system?

[ ] Windows	[**Terminate**]
[ ] Mac	[Continue]
[ ] I don't use a computer at all	[**Terminate**]

[**Recruiter**: Must use a Mac-based computer.]

3. Do you have access to high speed internet?

[ ] Yes	[Continue]
[ ] No	[**Terminate**]

[**Recruiter**: Must have access to high-speed internet because the test will be conducted and recorded via skype.]

4. Do you have a private working space with no distractions (e.g.: child, pets, loud noises)?

[ ] Yes	[Continue]
[ ] No	[**Terminate**]

[**Recruiter**: Participants must not be distracted during the experiment.]

Appendix 3

Questionnaire

1. Do you agree with this statement?

I know how to setup side-chain compression in my DAW.

[ ] Agree

[ ] Disagree

2. Do you agree with this statement?

I have applied side-chain compression in my previous project(s).

[ ] Agree

[ ] Disagree

(**Moderator** : Skip question 3 if the participant disagree.)

3. In your previous project(s), did you often use side-chain compression for corrective purposes (e.g., reducing masking) or creative purposes (e.g., pumping effect)?

[ ] Corrective

[ ] Creative

[ ] Both

4. For this experiment, did you use the nonlinear mixing system to create pumping effect or to reduce masking?

[ ] Create pumping effect

[ ] Reduce masking

[ ] Both

5. Do you agree with this statement?

There is an advantage / are advantages of using this system.

[ ] Agree

[ ] Disagree

(**Moderator** : Skip question 6 if the participant disagree.)

6. What is the advantage / are the advantages?

7. Do you agree with this statement?

I discovered a mixing method / technique that is unique to this system.

[ ] Agree

[ ] Disagree

(**Moderator** : Skip question 8 if the participant disagrees.)

8. What is the method / technique?

Appendix 4

Answer Transcript

Participant 1

1) Agree
2) Agree
3) Both
4) Corrective
5) Agree
6) The advantage is you can use more than two tracks for compression so for example you can set it like… Usually on the normal DAW that I use you only can set it to one track. This system uses multiple, so I think that's a really good system to be used for future project.
7) Agree
8) When I first use it like I realize that we can change the tone (of the mix) using this compressor. So, it's not only based on volume or the power, but we can also change the tone. This what I've learned.

   So, what do you mean by tone can you explain more?

   The quality of the sound. For example, I use one of your sample loops which is the kick drum. It's like an EQ but it's not really like an EQ… But there are some differences I can hear.

   So, you can't do this in DAW right

   Yeah, I think different types of of compressor produces different types of tone quality.

Participant 2

1) Agree
2) Agree
3) Creative
4) Corrective
5) Agree

6) For me the advantage of this system is I take less time to think, or I take less time to calibrate my system so it will be easier for the nonlinear mixing system because it does the thinking for me in a sense.
7) Disagree
8) Skipped

Participant 3

1) Agree
2) Disagree
3) Skipped
4) Creative
5) Agree
6) Normally when we use compression it's it depends on the sound, and it depends on the instrument. But I do believe for some of the creative purposes, especially for the current trend in in making music, this is the best solution that we have, to create pumping effect. Nowadays, people are more into this EDM music, and using a normal compression you can’t have this kind of effect. For the idea that you are showing me, I do like the idea where you have some assistance, or a system to calibrate or do some algorithm on applying sidechain compression.
   So, let me see if I got this correctly. So basically, what you meant is by having an algorithm that already have their side chain compression set up and then easily triggered by using button, it makes it easier. Is that what you mean?
   Yes correct.
7) Disagree (still looking into it).
8) Skipped.

Participant 4

1) Agree
2) Agree
3) Both
4) Both

5) Agree
6) First, if you need a sidechain compression it will be just very fast. Routing is already done. You don't have to create your own routing or doing sidechain compression for both previously mentioned purposes, creative pumping effect or reducing masking effect. Because they all are sidechaining together, I actually made some very interesting rhythmic effect that I couldn't achieve by only sidechaining one track or using traditional sidechaining method. So, I really enjoyed using that application for creative purpose. For the masking effect, I noticed that it would definitely have effects but at the same time I found it a little bit difficult to control. I definitely find it a little bit different on trying to balance the tracks. But I haven't conducted a very detailed experiment on that. Based on the knowledge that I have about compression; this system will definitely help to reduce masking. But it would also have disadvantages because using sidechaining to reduce the volume… the pumping effect is actually causing a clarity. You're trying to dig up the buried sound from the other tracks by using volume reduction… automated volume reduction. So, for that purpose it's why I can definitely write a paper on that, I’m not sure whether it will good or not but that's my observation for now.
7) Agree
8) I created a like a different pattern or a rhythmic pattern. But it's based on the sidechaining method. That means that it's a different technique. I sidechained everything. So, I played the pad... I used your material… the pad is quite consistent. I can hear a lot of rhythmic patterns on the volume reductions and it's interesting. I just think of one way that you probably could achieve the same result. For example, I load all the audios into separate tracks in the DAW. Then I create an auxiliary track or a bus. After that I send every single track to the bus. But the bus’s output it's selected to no output or sent to a to a compressor that allows sidechaining. The sidechain source of the compression is the auxiliary track or the bus. It could probably achieve the same result. But this system saves quite a lot of time. I haven’t tried it myself but in my mind, I’m thinking maybe it could.

Participant 5

1) Agree
2) Agree
3) Both
4) Corrective
5) Agree
6) It is easier for me to notice the pumping effect which makes enables me to define any dynamic changes or ‘movements’ in the arrangement. Especially when I can just change the value of the attack and threshold just to do so. It is also easier for me to notice which audio signal is sidechained indicated by the sidechain button at the bottom of the track. The set up is also very simple whereby I can just use a button to engage sidechain compression compared to the complicated process in DAWs such as Logic Pro X.
7) Agree
8) It is easier for me to do gain staging using the input gain, the compressor and the make up gain. It also changed the timbre of the samples that I upload in the system. For example, while mixing hip hop and trap tracks, when I used 808-kick drum the track did not peak. Additionally, the sound is punchier. Usually when I used stock Logic Pro X plugin, the sound tends to get buried and muddy. I had to do additional processing prior to compressing to get the sound I want. Your system gave the 808-kick clarity and punch which is exactly what I am looking for. Even when I use a lot of low-frequency elements in the mix which is typical in trap and hip hop, the sound is still clear, and you can hear each of the low-frequency elements individually.

Participant 6

1) Agree
2) Agree
3) Both
4) Corrective
5) Agree

6) Personally, I would use it to… Hmmm… If I'm lazy then I would use it to just… Bring out elements that are jumping out a bit that needs to be um… How do I say it… I think from my last use, whatever that's like louder, it will automatically bring everything else down. So, whenever there's a lead melody or lead instrument in there, it's easy for me to just let it just sidechain on its own and then I don't have to touch (adjust) anything anymore.
7) Agree
8) I think it's kind of is pretty much the same as my previous answer. It's like it is an automatic lazy mode, lazy sidechain mode. Like it is kind of just a… What's the word… It just detects which element is like louder, and then it highlights that and then I'm like "Okay, I don't have to touch anything…

Participant 7

1) Agree
2) Disagree
3) Skipped
4) Corrective
5) Agree
6) So, from what I understand the advantage is you can create mutual sidechaining of two different tracks from example drums to bass and vice versa. I have tried doing this in Logic Pro X and what you have mentioned about this not being able to be created in Logic is correct. There will be some sort of unwanted effect, so it can only go one way. If bass is sidechained to the drums, then that's it you can't do it the other way around. Yes, you can do the routing, but the outcome is not desirable. Maybe the sound is too compressed, I am not sure. I need to do more experiment on it. But using your system if I turned on the sidechains it will still sound good. It is usable in a mix.
7) Agree
8) So again, it's like what I have explained earlier. I can use this massively mutual sidechain in my mix because it is not possible in Logic. If I only use one track as the dominant, the output is not the same as if I use more than one. So, the approach is different.

Participant 8

1) Agree
2) Agree
3) Both
4) Creative
5) Agree
6) It's easier using this system. Everything is set up into an interface just like a DAW where you can see the tracks that you've loaded for the sidechain (indicated by the button) rather than having to click on the bus. Plus, you can also see the peak level on the meter for all the tracks together.
7) Agree
8) This is one unique way of mixing, one unique system. For me, what is unique… First of all is the interface. It's easy for you to see and control everything. And then its easy for me to set up the amount of compression from track to track. And I think you can sidechain a whole bunch of tracks into once sidechain, if I'm not mistaken, through this system. Everything is linked from one track to another. So, I see it is unique.

Participant 9

1) Agree
2) Agree
3) Creative
4) Both
5) Agree
6) I guess because it it's sort of like a compounding compression, so if for example instead of using sidechain compassion for sort of subtractive purposes you can use it as a way to sort of bring up certain elements that you want to bring up in in terms of like a mixing terminology.

   I see so does it mean it makes certain elements that you want into mix clearer or to bring it up more in the mix.

Yes. Also, for example, so what I did was I tried different audio samples. So basically, some of the audio samples I did was was stems. So just a basic kick loop, chords et cetera. That's just a pure normal mix. Another one that I did was I’ve imported different parts of my drums. Snare on one track, kick on one track, and then cymbals on one track. So, I think where this one really shines is, for example, you can use it as an over-the-top (OTT) in a way for like the crash cymbal. Instead of like automating everything, you pump down the hit of the crash but then tail is unaffected. I guess like a transient shaper, you can use it like a transient shaper.

7) Agree
8) I'd say one area where this might be interesting to look at it is in terms of like mixing drums or transient shaping. I'm sure if we have more time, we can find other ways yeah of using it.

Participant 10

1) Agree
2) Agree
3) Both
4) Corrective
5) Agree
6) It's a new approach to sidechaining. Because the way we always use it, it was always for a certain purpose, which is reducing masking and pumping effect. For me it broadens up your options on what you can do with the system.
7) Agree
8) First-hand experience of using it… Well, I'm not used to having multiple sidechains you see. So, when I first use it, the effect that I get from it it's different. So, I started exploring and exploring. I don't know how to put it in words. It's unique I don't know how to say it dude. When I was using DAW before this it's only a two-way thingy… You know. So, it only effects two tracks. Right now when you have eight simultaneously being sidechain… So, I'm discovering new ways that when one is sidechained, how it affects the other

instruments. And finding a balance into it. I guess that's a new way of looking at things and hearing it. So, I guess that's the plus experience that I got from it. I am so used to using sidechain, for example let's say you have a kick and a bass. And you want to rid of the muddiness especially when the kick is playing. I'm so used to only using it for that, not so much to get the pumping effect. The type of processing that I use in the music that I produce is mostly involving just two elements with one controlling another. So, what I experienced using your system it blew my mind to see how it works. It took a while to get used to it. So yeah, we discover new things… So, in one of my productions I have two types of drums and two types of bass going on simultaneously. If I were to use this system, I would apply it in that production. Because since I have two drums and two basses, I have to apply sidechain separately from one drum to one bass and then on the other section I have to apply a different sidechain. To save time maybe, you can put all the tracks one section after another and then sidechain it in one go, so I don't have to do it twice. That's how I would use it if I was using multi-timbral basses in my song. This new system of yours will broaden up new ways of production because now producers won't be scared of using more low-frequency elements in their song. If I have this system, I will just put everything inside it, mix it there, bounce it out and then put it in my DAW.

And having buttons to trigger sidechain is a simpler way of doing it rather than having to route everything one by one.

Participant 11

1) Agree
2) Agree
3) Both
4) Creative
5) Agree
6) One of the advantages when I'm using the same system is I can simply choose which one of the instruments in the music that will be the dominant. The second advantage is that I can prevent my music from clipping. This is because I can reduce the volume of any

specific instrument in the system (including the dominant) because sometimes the problem when you're mixing is you often get a clipping.

7) Agree
8) Yes, because most of the available system right now only allows sidechaining only one track or one channel. But, in this system you can a combine two or three tracks together to affect one track or more. So, I think this is an advantage during mixing because it opens many doors for creativity. So, have you found any new methods specifically that can only use in this application? Is it just setting up one or two instruments controlling the other? When I'm playing arpeggios, for example, I can set up three tracks and each track is playing different notes. For example, 1-3-5 1-3-3 or 1-3-7. So, I can apply this system to create new atmosphere or sound.

Participant 12

1) Agree
2) Agree
3) Corrective
4) Both
5) Agree
6) Let's say if I want to reduce masking. I can use one of the main tracks, for example, kick drum to reduce the the volume of the bass whenever it clashes. Kick drum will cut through the mix. I can do the same thing on the synth pads playing sustained chords. I will do the same thing as what I do with the bass. With your system I can use kick drum track to sidechain it to the bass and the chords. So, its easier you don't have to do route all the sidechains on your DAW, with your system it is already routed.
7) Agree
8) Before this I thought that I only sidechain one track to another. So, with this system I found a new method in which I combined a few tracks maybe two or more and use it as the source track for sidechaining.

9) **Participant 13**

1) Agree

2) Agree

3) Both

4) Both

5) Agree

6) I think this would be very strong for music that's ambient in nature. So, let's say you have this droning effect that you would want to… This is just a very basic example. You have two drones, and you want a specific drone to still be slightly louder than the other ones. Depending on whether there are other instruments as well, such as a snare, you could then use that snare input to be the deciding factor in controlling the compression or the output level of that of that drone. So, the thing is… The interesting part is that you don't know what will happen… That would be the take there. So, it sometimes might not be an ideal result but there's also a 50% chance of it being a very very interesting effect.

7) Agree

8) The mixing method I've found is that this system uses the the input level of each individual track. So, if you want to play around with, let's say for me this is the track that is most prominent, you can actually decrease levels of all the other tracks simultaneously. To what exact number I'm not too sure, but as long as you boost the input level of one track all the other tracks are also affected. So, in a way it could be a very good method of, you know, performance type of setting. Let's say in the mix you want to particularly boosts the signal of a snare drum and you want all the other tracks all the other instruments to also be quiet without you having to do it 1 by 1. So, I can see it in a performance method because if you were to do this in the mixing where, yes you could just reduce all the other levels individually but that will take some time. But in this method, you just need to raise the input level of one track and everything else is affected. So, in a way it makes it easier. And also, there's a lot of lot more creative aspects to it as well. You can use it to accentuate some parts of the song you can also use it to create a pumping sort of effect as well. So yeah, a

lot of ways to mess around with this plugin. Just that probably with other people they might find a different way to do it.

Participant 14

1) Agree
2) Agree
3) Corrective
4) Creative
5) Agree
6) The control surface allows for tactile control and enables creative possibilities during mixing. This is also good for performance control surface and beat making for EDM. Just load some loops in the system and record. It can also be used for song writing and arrangement tool. So, you can just throw random stuff in there then use the sidechain the massively mutual sidechain compression to make some surprises so to speak. The last practical use in it came up it's just it's for radio. If they have multiple audio sources, you can use it as an auto mix itself. Because I figured… I read the documents that you included the threshold of the signal depends on the income level of other signals. So, if you want to control a signal and you can just utilize that your own way to… for you know… how do you say… automatic radio mix.
7) Agree
8) Its not normal though I have to put it first. Because most mixing engineers would use this sort of recursive system to use sidechain compression, but it doesn't do what this one does. And it's really insane actually. It's just the system itself is new… for me. Maybe I'm just idiot… I haven't discovered this sort of thing before. But yeah, it's definitely new. The recursive sidechain gives the chance from having one signal react to control this these two. For instance, like kick to bass, or kick to something like guitar or whatever. Or vocals… to the whole instrument and bus sort of thing.

Participant 15

1) Agree
2) Agree
3) Corrective
4) Corrective
5) Agree
6) For me it is a new approach, usually the one that I used in my DAW is one-to-one compression for example snare and bass drum. I've never encountered a massively mutual compression system such as this system. Furthermore, this system is a standalone application (not a plugin). It's like its own mixer without needing to load it in a DAW. When I run the system, the routing is already set up. And you can clearly hear how the system affect the tracks. Masking is reduced and it sounds more detailed. For example, the bass and the kick. The kick sounds more energetic because the bass guitar is slightly ducked. The best part is it also slightly ducks the other elements as well without me having to route it.
7) Disagree. (I haven't explored it enough).
8) Skipped

Resource List

Bibliography

Houtsma, A.J.M. 1997. “Pitch and Timbre: Definition, Meaning and Use.” *Journal of New Music Research vol. 26 issue 2* 26 (2): 104-115.

Zagorsky-Thomas, Simon. 2007. “The Study of Groove.” *Ethnomusicology Forum* 16 (2): 327-335.

Danielsen, Anne. 2006. “Interaction of Rhythm and Sound in Contemporary Dance Music.” *Art of Record Production.* 9 September. Accessed January 14, 2020. https://www.artofrecordproduction.com/aorpjoom/symposiums/18-arp-2006/88-danielsen-2006.

Zeiner-Henriksen, Hans T. 2006. “The Most Significant Beat: A Comparatice Study of Changes in Bass Drum Sounds from 70s Disco to Electronic Dance Music of the 1980s and 1990s.” *Art of Record Production.* 9 September. Accessed January 14, 2020. http://www.artofrecordproduction.com/aorpjoom/arp-conferences/arp-2006/18-arp-2006/102-zeiner-henriksen-2006.

Hodgson, Jay. 2011. “Lateral Dynamics Processing in Experimental Hip Hop: Flying Lotus, Madlib, Oh No, J-Dilla and Prefuse 73.” *Journal on the Art of Record Production.* July. Accessed January 14, 2020. http://www.arpjournal.com/asarpwp/lateral-dynamics-processing-in-experimental-hip-hop-flying-lotus-madlib-oh-no-j-dilla-and-prefuse-73/.

Zeiner-Henriksen, Hans T. 2012. “Moved by the Groove: Bass Drum Sounds and Body Movements in Electronic Dance Music.” In *Musical Rhythm in the Age of Digital Reproduction*, edited by Anne Danielsen, 121-139. Surrey: Ashgate Publishing Ltd.

Lacasse, Serge. 2012. “Slave to the Supradiagetic Rhythm: A Microrhythmic Analysis of Creaky Voice in Sia's 'Breathe Me'.” In *Musican Rhythm in the Age of Digital Reproduction*, edited by Anne Danielsen, 141-155. Surrey: Ashgate Publishing Ltd.

Solberg, Ragnhild Torvanger, and Alexander Refsum Jensenius. 2016. “Pleasurable and Intersubjectively Embodied Experiences of Electronic Dance Music.” *Empirical Musicology Review* 11 (3-4): 301-319.

Brown, Judith C. 1991. “Calculation of a Constant Q Spectral Transform.” *Journal of the Acoustical Society of America* 89 (1): 425-434.

Schörkhuber, Christian, and Anssi Klapuri. 2010. “Constant-Q Transform Toolbox for Music Processing.” *7th Sound and Music Computing Conference.* Barcelona. 3-64.

Mathworks. n.d. “Nonstationary Gabor Frames and the Constant-Q Transform.” *Mathworks.* Accessed February 2020, 9. https://uk.mathworks.com/help/wavelet/gs/non-stationary-gabor-frames.html.

Danielsen, Anne. 2019. “Where is the Beat in That Note? Effects of Attack, Duration, and Frequency on the Perceived Timing of Musical and Quasi Musical Sounds.” *Journal of Experimental Psychology: Human Perception and Performance* 45 (3): 402-418.

White, Paul, Hugh Robjohns, and Dave Lockwood. 2013. “Drums.” In *The Studio SOS Book: Solutions and Techniques for the Project Recording Studio*, 176. Massachussets, Burlington: Focal Press.

Lartillot, Olivier, Petri Toivianen, Pasi Saari, and Tuomas Eerola. n.d. "MIRtoolbox." *University of Jyväskylä.* Accessed September 29, 2020. https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox.

Hallum, Richard. 2016. "Measurement of Harmonic Distortion Characteristics in Logic Pro X Compressor Plugin." *Researchgate.* July. Accessed September 29, 2020. http://doi.org/10.13140/RG.2.2.33180.49288.

Rogerson, Ben. 2019. "How Stardust Made Music Sounds Better With You... And Why It Was Their Only Release." *Music Radar.* June. Accessed December 19, 2019. https://www.musicradar.com/news/how-stardust-made-music-sounds-better-with-you-and-why-it-was-their-only-release.

Randolph, Robert. 2019. "DAW v DAW - Part 5: Plugin Automation." *AdmiralBumblebee.* 22 June. Accessed August 14, 2020. https://www.admiralbumblebee.com/music/2019/06/22/Daw-V-Daw-Automation-Part-4.html.

Vandemast-Bell, Paul. 2013. "Rethinking Live Electronic Music: A DJ Perspective." *Contemporary Music Review* 32 (2-03): 239-248.

Rietveld, Hillegonda C. 2017. "Authenticity and Liveliness in Digital DJ Performance." In *Musicians and their Audiences: Performance, Speech and Mediation*, by Ioannis Tsioulakis and Elina Hytönen-Ng, 123-133. New York: Routledge.

Apple Inc. n.d. "Logic Pro Klopfgeist." *Apple Support.* Accessed August 2020, 15. https://support.apple.com/en-gb/guide/logicpro/lgcp67edcb1f/mac.

vangarecord. 2014. *How to Sidechain by Klopfgeist on Logic Pro X.* 13 December. Accessed August 15, 2020. https://www.youtube.com/watch?v=texH6iJFRG0.

imamusicmogul. 2016. *How to Sidechain Like a Boss Using a Ghost Trigger Track.* 24 September. Accessed August 15, 2020. https://www.youtube.com/watch?v=tN6Ql7coRGU.

Sterling, Christopher H., and Cary O'Dell. 2010. "Audio Processing in the Audio Chain." In *The Concise Encyclopedia of American Radio*, edited by Christopher H. Sterling, 77-78. New York: Routledge.

Somich, Jim, and Barry Mishkind. 2009. "A History of Audio Processing Part 1." *The Broadcasters' Desktop Resource.* Edited by Barry Mishkin. September. Accessed March 10, 2020. https://www.thebdr.net/articles/audio/proc/ProcHist.pdf.

Collett, John Peter. 2003. "The History of Electronics." In *Companion to Science in the Twentieth Century*, edited by John Krige and Dominique Pestre, 254. London: Routledge.

Somich, Jim, and Barry Mishkind. 2009. "A History of Audio Processing Part 2 - Automatic Processing Starts." *The Broadcasters' Desktop Resource.* September. Accessed March 10, 2020. https://www.thebdr.net/a-history-of-audio-processing-part-2-automatic-processing-starts/.

Hood, John Linsley. 1997. "Valves or Vacuum Tubes." In *Valve and Transistor Audio Amplifiers*, by John Linsley Hood, 1. Oxford: Newnes.

Bieger, Hannes. 2016. "Fairchild 660 and 670 Legendary Compressor-Limiter." *Sound on Sound Review.* May. Accessed August 24, 2017. https://www.soundonsound.com/reviews/fairchild-660-670.

Snoman, Rick. 2009. "Compression." In *Dance Music Manual: Tools, Toys and Techniques*, 108. Oxford: Elsevier Ltd.

Siekmeier, D. 1962. “An Apparatus for the Real-Time Transmission of Handwriting and Map Information to Remote Displays.” *Defence Tehcnical Information Centre.* January. Accessed April 3, 2020. https://apps.dtic.mil/docs/citations/AD0269991.

Case, Alexander U. 2007. “Optical Compressors.” In *Sound FX: Unlocking the Creative Potential of Recording Studio Effects*, by Alexander U. Case, 138. Burlington: Focal Press.

Shanks, Will. 2003. “Compression Obsession: The Amazing Release Character of the LA-2A.” *Universal Audio Webzine.* 3 June. Accessed August 15, 2017. https://www.uaudio.com/webzine/2003/june/text/content4.html.

Fuston, Lynn. 2012. “Universal Audio's Classic 1176 Compressor - A History.” *Universal Audio Blog.* 15 May. Accessed April 6, 2020. http://www.uaudio.com/blog/analog-obsession-1176-history/.

Barbour, Eric. 1999. “The Cool Sound of Tubes.” *IEEE Spectrum.* 4 Jan. Accessed April 4, 2020. https://spectrum.ieee.org/consumer-electronics/audiovideo/the-cool-sound-of-tubes.

Hamm, Russell O. 1973. “Tubes Versus Transistors - Is There an Audible Difference?” *Journal of the Audio Engineering Society* 21 (4).

Hewlett-Packard. n.d. “HP Virtual Museum: HP-35 Handheld Scientific Calculator, 1972.” *Hewlett-Packard.* Accessed March 12, 2020. http://www.hp.com/hpinfo/abouthp/histnfacts/museum/personalsystems/0023/.

Engineering and Technology History Wiki. n.d. “Milestones: Development of the HP-35, the First Handheld Scientific Calculator, 1972.” *Engineering and Technology History Wiki.* Accessed March 12, 2020. http://ethw.org/Milestones:Development_of_the_HP-35,_the_First_Handheld_Scientific_Calculator,_1972.

Ballou, Glen. 2015. “Integrated Circuits for Audio Applications.” In *Handbook for Sound Engineers*, edited by Glen Ballou, 341. Burlington: Focal Press.

Giannoulis, Dimitrios, Michael Massberg, and Joshua D. Reiss. 2012. “Digital Dynamic Range Compressor Design - A Tutorial and Analysis.” *Jurnal of the Audio Engineering Society* 60 (6): 399-408.

Verfaille, Vincent, Daniel Arfib, F. Keiler, A. von dem Knesebeck, and Udo Zölzer. 2011. “Adaptive Digital Audio Effects.” In *DAFX: Digital Audio Effects*, edited by Udo Zölzer, 324. West Sussex: John Wiley & Sons Ltd.

Reiss, Joshua D., and Øyvind Brandtsegg. 2018. “Applications of Cross-Adaptive Audio Effects: Automatic Mixing, Live Performance and Everything in Between.” *Frontiers in Digital Humanities* 5 (17): 2-4.

Hodgson, Jay. 2010. “Ducking.” In *Understanding Records: A Field Guide to Recording Practice*, 95. New York: The Continuum International Publishing Group Inc.

Wright, Tim. 2013. “Folio of Compositions.” *White Rose eTheses Online.* 19 November. Accessed September 29, 2020. http://etheses.whiterose.ac.uk/4612/.

ReaperBlog. 2016. “Delay Feedback Effects.” *ReaperBlog.* July. Accessed July 11, 2020. https://reaperblog.net/2016/07/delay-feedback-effects/.

Stillwell, Thomas Scott. 2006. “JS Major Tom Compressor Source Code.” Reaper Digital Audio Workstation.

Taylor, Gregory. 2018. “gen~ for Beginners, Part 1: A Place to Start.” *Cycling '74 Tutorial.* 28 February. Accessed April 20, 2020. https://cycling74.com/tutorials/gen~-for-beginners-part-1-a-place-to-start.

Bohn, Dennis. 2005. “Linkwitz-Riley Crossovers: A Primer.” *RaneNote.* October. Accessed April 2020, 2020. https://www.ranecommercial.com/legacy/note160.html.

Marrington, Mark. 2017. “Mixing Metaphors: Aesthetics, Mediation and the Rhetoric of Sound Mixing.” In *Mixing Music (Perspectives on Music Production)*, edited by Russ Hepworth-Sawyer and Jay Hodgson, 199-215. New York: Routledge.

Apple Inc. n.d. “Plug-ins and Sounds.” *Logic Pro X.* Accessed July 7, 2020. https://www.apple.com/uk/logic-pro/plugins-and-sounds/.

—. n.d. “Logic Pro X: Stereo Spread.” *Logic Pro X User Guide.* Accessed July 18, 2020. https://support.apple.com/kb/PH27281?locale=en_US&viewlocale=en_US.

—. n.d. “Logic Pro Klopfgeist.” *Logic Pro X User Guide.* Accessed August 15, 2020. https://support.apple.com/en-gb/guide/logicpro/lgcp67edcb1f/mac.

—. n.d. “Logic Pro X: Mixer Interface.” *Logic Pro X User Guide.* Accessed April 14, 2020. https://support.apple.com/kb/PH24681?locale=en_US&viewlocale=en_US.

—. n.d. “Logic Pro X: Gain Plug-In.” *Logic Pro X User Guide.* Accessed August 14, 2020. https://support.apple.com/kb/PH27564?locale=en_US&viewlocale=en_US.

Hove, Michael J., Céline Marie, Ian C. Bruce, and Laurel J. Trainor. 2014. “Superior Time Perception for Lower Music Pitch Explains Why Bass-Ranged Instruments Lay Down Musical Rhythms.” *Proceedings of the National Academy of Sciences of the United States of America.* 10383-10388.

Wagenaars, Wil M., Adrianus J. Houtsma, and Ruud A. van Lieshout. 1986. “Subjective Evaluation of Dynamic Compression in Music.” *Journal of the Audio Engineering Society* 34: 10-18.

Madison, Guy, and George Sioros. 2014. “What Musicians do to Induce the Sensation of Groove in Simple and Complex Melodies, and How Listeners Perceive It.” *Frontiers in Psychology* 5 (894): 1-14.

Sioros, George, Marius Miron, Matthew Davies, Fabien Gouyon, and Guy Madison. 2014. “Syncopation Creates the Sensation of Groove in Synthesized Music Examples.” *Frontiers in Psychology* 5 (1036): 1-10.

Fikentscher, Kai. 2000. In *You Better Work!: Underground Dance Music in New York*, 88. Connecticut: Wesleyan University Press.

Stardust. 1998. *Music Sounds Better with You.* Comps. Alan 'Braxe' Quême, Benjamin 'Diamond' Cohen and Thomas Bangalter.

Lotus, Flying. 2007. “Tea Leaf Dancers.” *Reset.* Comp. Flying Lotus.

Timberlake, Justin. 2016. *Can't Stop the Feeling.* Comp. Justin Timberlake.

Morse, Janice M. 2012. “The Implications of Interview Type and Structure in Mixed-Method Designs.” In *The SAGE Handbook of Interview Research*, edited by Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti and Karyn D. McKinney, 194. Los Angeles: SAGE Publications, Inc.

Creswell, John W., and J. David Creswell. 2018. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* 5th Edition. Los Angeles: SAGE Publications, Inc.

Beitin, Ben K. 2012. “Interview and Sampling: How Many and Whom.” In *The SAGE Handbook of Interview Research*, edited by Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti and Karyn D. McKinney, 244. Los Angeles: SAGE Publications, Inc.

Johnson, John M., and Timothy Rowlands. 2012. “The Interpersonal Dynamics of In-Depth Interviewing.” In *The SAGE Handbook of Interview Research*, edited by Jaber F. Gubrium, James A. Holstein, Amir B. Marvasti and Karyn D. McKinney, 108. Los Angeles: SAGE Publications, Inc.

Solberg, Ragnhild Torvanger, and Alexander Refsum Jensenius. 2017. “Group Behaviour and Interpersonal Synchronization to Electronic Dance Music.” *Musicæ Scientæ* 23 (1): 111-134.