A Quantitative Evaluation of Signal Masking in Summed and Compressed Audio

Toulson, R., Paterson, J. and Campbell, W.

This is a pre-publication version of a book chapter published in Innovation In Music, Future Technology Press, Shoreham-by-Sea, UK. pp. 20-31 .

KES Transactions on Innovation in Music: Vol 1 No 1
Special Edition - Innovation in Music 2013 : pp.20-31 : Paper im13bk-002

A Quantitative Evaluation of Signal Masking in Summed and Compressed Audio

William Campbell[1], Justin Paterson[2], Dr. Rob Toulson[3]

[1,3] Anglia Ruskin University, East Road, Cambridge, CB1 1PT, UK

[2]London College of Music, University of West London, St. Mary's Rd., Ealing, London, W5 5RF, UK

1. Abstract

In music production, it is common practice to apply dynamic-range compression to audio signals. Traditionally, the operator's attention is drawn to the reduction in dynamic range and the sonic signature imposed by the envelope (settings) of the device, and the resulting distortions are familiar to studio practitioners. However, the non-linear characteristics of compression, combined with the interaction of these signals once summed, are likely to produce less familiar side effects, such as intermodulation distortion, manifesting itself as signal masking and other related artefacts.

Comparative quantitative analysis of compressed simple and compound signal structures shows the products of this distortion to be realignment of harmonic structure, reduction of spectral and temporal clarity, and rearrangement of dynamic variances related to the rhythmic structure of musical signals. Although the rearrangement of the dynamic variances is expected (in that the variances are reduced), what is less expected is the extent to which amplitudes of certain individual components of summed signals are attenuated, effectively precipitating signal masking. This research shows that decreasing the number of signals interacting with each other whilst applying an equivalent amount of compression can reduce the intermodulation distortion and therefore improve the overall signal quality of commercial music.

2. Introduction

This paper explores the non-linear properties of dynamic compression on simple signals and the effects that the observed nonlinearities have on musical signals. Here we will utilise extreme dynamic range compression applied to simple signals and also investigate its effects on musical signals. Such use of heavy compression (in the form of limiting) helps us to understand the associated nonlinearities by highlighting the artefacts that are present in all compressed signals, but perhaps less observable when lesser amounts of compression are used. We will also discuss early experimental findings in which compression outcomes with simple signals and processing chains are examined.

At this time, these experiments are restricted to the application of digital compression, and although the findings here may also hold true for hardware compression, it is not explicitly addressed.

3. Background and Related Work

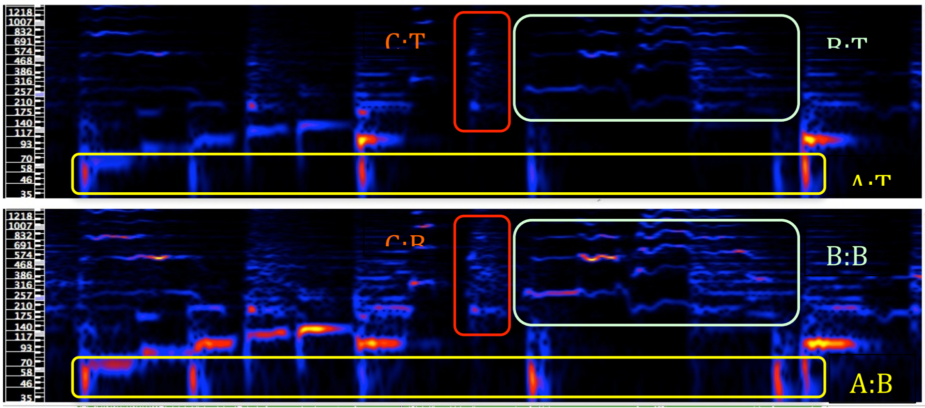
Over the history of record production, music has been gradually getting louder [1]. The perceived level or 'loudness' of a piece of music is very closely related to the RMS (root-mean-square) of the audio signal. In recent years there has been an increase in RMS values of music programmes, which is a result of the use of dynamic range compression on the audio signal. Application of compression to a dynamic signal reduces the part of the signal that rises in amplitude above a designated threshold, reducing the variance between the quietest and loudest portions of the piece. This reduction in amplitude variation allows the overall signal amplitude to be increased through application of makeup gain (increasing the overall signal amplitude to the maximum digital full-scale level before clipping) to increase the RMS level.

Compression is a non-linear process, and in some contexts is related to waveshaping [2]. A system is non-linier when a signal or signals input into some form of processing, has alterations to its spectral content. If the spectral content is unaltered, but the amplitude is altered, the system is linier. 'A non-liner process creates harmonic and inharmonic frequency components, not present in the original input signal [3].'

This research aims towards replicating the compression levels required to maximise the RMS of music programmes and increase loudness in a fashion comparable to modern popular music. Previous research has shown some modern popular music to have RMS levels (AES Standard) to be as high as –1.11 dBFS [4]. Achieving this level maximisation is best attained via limiting, an extreme form of compression [5,6].

It has been shown that heavy compression does not simply reduce the dynamic variance of music — it has a much more detailed effect on the audio signal. Figure 1 is a spectrogram of an excerpt of a musical passage recorded for the experimental purposes of this on-going PhD research. The top (suffix :T) signal is uncompressed whilst the bottom (suffix :B) is compressed using a limiter The peak signal level of the compressed signal vs. the uncompressed signal are equalised, effectively adding gain to the compressed signal.

Figure 1 Spectrogram showing uncompressed (top) signals and compressed (bottom) signals demonstrating that compression does not simply reduce differences in signal amplitude.



For instance, variations in amplitude of single instruments may be rearranged depending on the interaction with other instrument signals as indicated in regions 1A:T and 1A:B. This rearrangement of the amplitude variations can have the effect of altering the rhythmic structure, indicative of the ‘feel’ of the piece [7]. This rearrangement of amplitude can also alter the harmonic structure and thus the tonality of individual signals [4,8], observed in regions 1B:T and 1B:B.

Early reflections, reverberations and background noise of the recording environment are amplified by compression. Regions 1C:T and 1C:B in Figure 1 show a solitary snare drum signal. The spectrogram of the compressed phrase shows noise in the area of the snare drum, the early reflections, reverberations and background noise of the room being amplified by the compressor. The increase in amplitude of the early reflections and reverberation may give the impression of reduction in depth and width of the stereo spectrum, or possibly that of comb-filtering [9].

In psychoacoustic research to analyse the effect of compression when applied to spoken word, Stone et al. [10] found that two voices being summed then compressed introduced intermodulation between the two voices. Their findings suggest that a reduction of the number of intermodulating signals in a signal chain reduces inherent distortions. When we apply this theory to musical programmes, the dynamic response is improved [8].

In this paper we examine the implications of generating intermodulation distortions of signals. We also look at how the number of intermodulation artefacts can be reduced through realignment of compression in the signal chain. We will finally look at the implications of this realignment on real musical signals.

4. Procedure/Experimentation

Two experimental procedures are discussed here. The first set of experiments is implemented using different configurations of mixing and compression of sine waves within Matlab. The second set of experiments is executed in Pro Tools with musical signals as the subject and analysed using Sonic Visualiser [11]. The relationship between the two experiments will be addressed in the analysis following.

4.1 Matlab Method

The purpose of this set of experiments is to analyse the effects of applying compression to sine waves, generated within Matlab at sampling frequency of 44.1 Hz. Each of the final mixed signals are analysed via the Matlab FFT function with a Blackman window applied and plotted in the time vs. frequency domain. There are two parts to this set: Test 1 — two sine waves mixed then compressed; Test 2 — two sine waves compressed then mixed. The two tests with mixed signals are implemented with both harmonically related signals and inharmonically related signals.

In order to exaggerate the potential compression artefacts, each of the test scenarios utilise a look-ahead brick wall limiter published in DAFX [3]. The limiter's attack time is 0.3 ms, a release time of 0.01 ms and a look-ahead (delay) of 5 ms. The delay is necessary for the limiter to be fast acting. The threshold in each of the tests is executed at three different levels; 0.5 (20log(0.5)=–6 dB) light compression; 0.3 (20log(0.3)=–10 dB), medium compression; and 0.1 (20log(0.1)=–20 dB), heavy compression.

In Test 1 we analyse two Matlab-generated sine waves of one second duration, mixed together and then compressed. There are two parts to this test. The first part is using sine waves that are harmonically related, 10 Hz and 20 Hz. The second part is with two inharmonically related signals, 10 Hz and 25 Hz. Altering the frequency relationship of the signals could offer some insight into the role of inharmonicity in mixed and compressed signals. The 'base' frequency of the sine wave, be it 20Hz or 20kHz has proven irrelevant in these experiments. The use of ultra low frequencies simplifies the analysis process and articulation of the charted results. The interval separating the frequencies is what we are interested in here.

In Test 2 two sine waves of one second duration are generated, compressed then summed. Again, there are two parts to this test, with frequencies as above, but, here we can explore the implications of harmonicity and in-harmonicity in signals compressed then summed, both for comparison to one another and for comparison to the previous test in which the signals are summed prior to compression.

4.2 Pro Tools Method

The Pro Tools part of the experiments is to analyse the effects of applying compression to musical signals. The configuration of the compression is similar to the Matlab experiment in that the placement of compression along the signal chain is being tested. The signal path (channel to sub-mix buss to master buss) for each of the test scenarios remains the same regardless of where the compression is inserted, Figure 2. Past experiments have shown no difference in signals due to routing through sub-mixes or not. The sub-mixes are configured as drums and percussion buss, vocal buss, and other instrument buss.

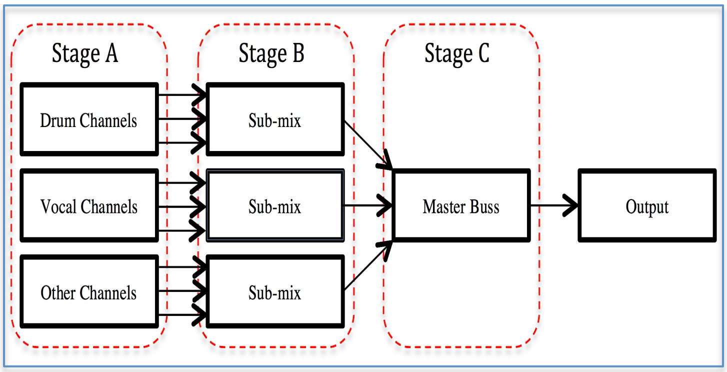


***Figure 2** Basic signal paths of all the Pro Tools experiments, demonstrating the three test stages for application of compression.*

Here we compare four different configurations; Test 1 — signals mixed with no compression applied, a control for comparison to subsequent configurations; Test 2 — compression applied at stage A in Figure 2 to the individual components of the mix (each channel) ; Test 3 — compression applied at stage B in Figure 2 to sub-mixes (stems); Test 4 — compression applied at stage C in Figure 2, after all signals are mixed, in other words, master buss compression.

The compression used for this part of the experiment is the Pro Tools native Dyn3 compressor/limiter with a 100:1 ratio, 10 μs attack time, 5 ms release time and –21 dB threshold. The Dyn3 compressor/limiter utilises look-ahead functionality [12], though the manufacturer does not specify the signal delay time. The parameters of the attack and release times are set to the fastest possible setting. The ratio is the highest available and the threshold is set to the approximate RMS level at the master buss.

5. Research Analysis

5.1. Matlab

In the first part of Test 1, results shown in Figure 3A and 3B, we compress two sine waves that are harmonically related; 10 Hz and 20 Hz. As the threshold is decreased the amount of compression increases, so too do the harmonically related distortion components. It is interesting to note that the increase in compression here has the effect of reducing the amplitude of the fundamental of the 10 Hz signal. It may be worth noting that ramping the amplitude of the signal does not add distortions to the signal, as it is a linear process. Using such a ramped signal allows us to observe the effect the onset of compression has on the signal envelope. However, altering the signal envelope using compression does add distortion components, as any non-linear process.

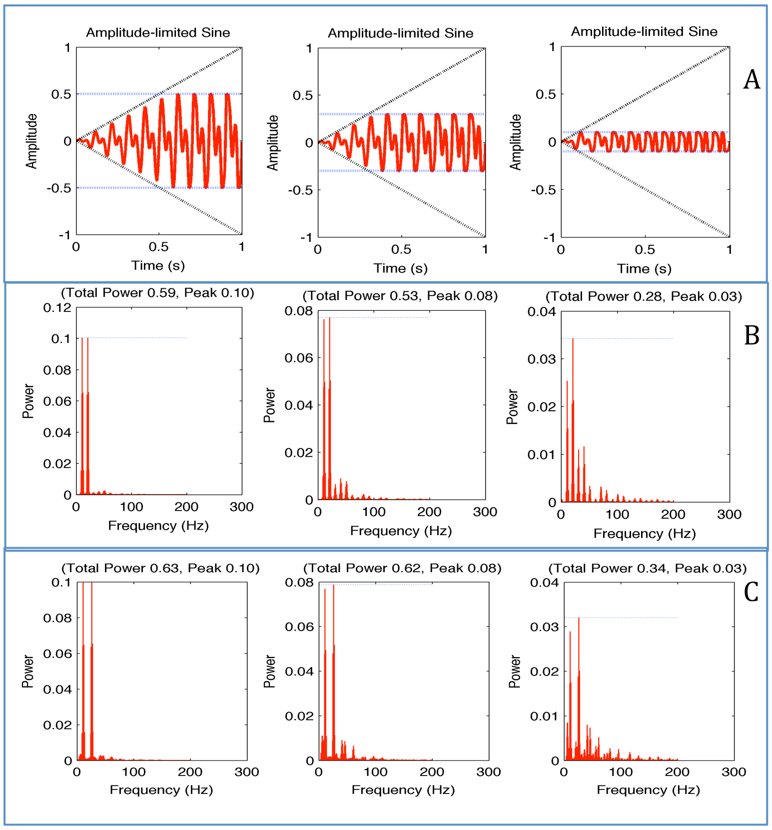


Figure 3 *Test 1.1 Signal analysis of two summed sine waves showing incremental amounts of compression. The top set of windows (A) showing the compressed signal envelops; (B) the corresponding FFT of the two harmonically related compressed signals; (C) the FFT of the two harmonically unrelated compressed signals.*

If we change the second signal (results shown Figure 3C) to an inharmonically related signal, again we observe distortions that are related to the two signals, but in addition can see sum and difference frequencies; intermodulation distortion. This type of distortion is not pleasant to listen to as it is not harmonically related to the original signals and is therefore perceived as more dissonant. Again we can see the amplitude reduction of the 10 Hz signal. Although this reduction is not as profound as when the signals are harmonically related, it is still present.

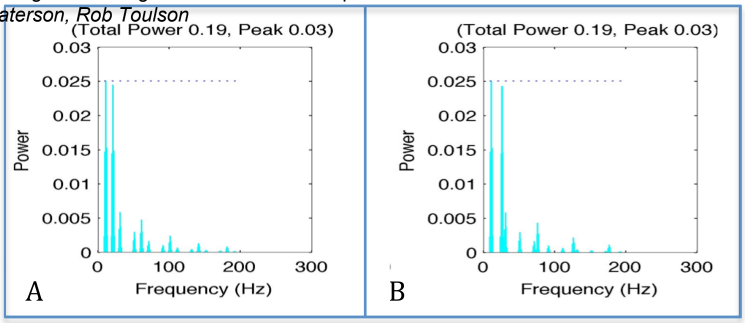


Figure 5 *Test 2 FFT analysis of two harmonically related signals) compressed and then summed (A) and two harmonically unrelated signals) compressed and then summed (B).*

In Test 2, results shown in Figure 5, the signals are compressed before mixing the signals together. Again we are using signals that are harmonically related in the first part of this test (Figure 5A) and inharmonically related signals in the second (Figure 5B). Here we observe that mixing the signals after applying compression significantly reduces the distortion artefacts, regardless of the signals being harmonically related or not. It is noteworthy that there appears to be cancellation of some of the distortion components in both the harmonically related and inharmonically related windows. This may be to do with the phase relationship of the harmonic components added by compression, because a number of sum and difference values calculate as integer harmonics of the two source signals when the two source signals are themselves harmonically related.

5.2. Pro Tools

Rather than exploring the effects of compression on a full musical passage, here we are making observations based on measurements in 1 second of a musical passage, at 0.005 ms. intervals. We are interested in the micro-dynamics of compression of a single note, concurrently in five different instruments: vocal, guitar, bass, conga and bass drum.

Figure 6 quantifies the uncompressed signals, plotting only the dominant frequency of each instrument. Note that each measurement clearly shows an onset period of each signal envelope. Once each instrument reaches a peak level, they have a short sustained section followed by a release section where the envelope has a decline in amplitude. Regarding the guitar, vocal and bass, the release period is not so obvious as it is constrained by the 1 s measurement window. The bass drum rises quickly and releases quickly to another sustained part of the envelope, a typical envelope of a percussive instrument. The envelope of the conga is notable, in that it rises and falls rapidly in an oscillating fashion, whilst still maintaining an overall percussive envelope (as indicated by the added grey trend line), likely displaying the low frequency components of this particular sound. We also observe that the guitar is the dominant instrument overall in this one second window, followed by the bass. The bass drum has the second highest amplitude initially during its onset phase, followed by the vocals and then the conga.

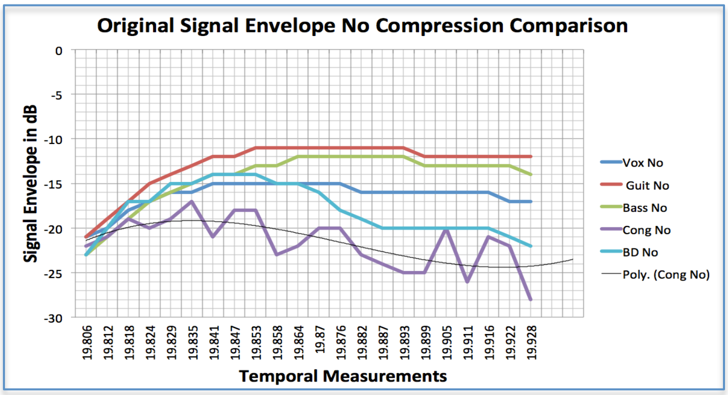


Figure 6 *Level measurements of 1 s excerpts of the uncompressed signals under test.*

However, the amplitude of each of the instruments is not really an indication of the overall perceived loudness because their holistic spectral content is also a factor in addition to their temporal placement.

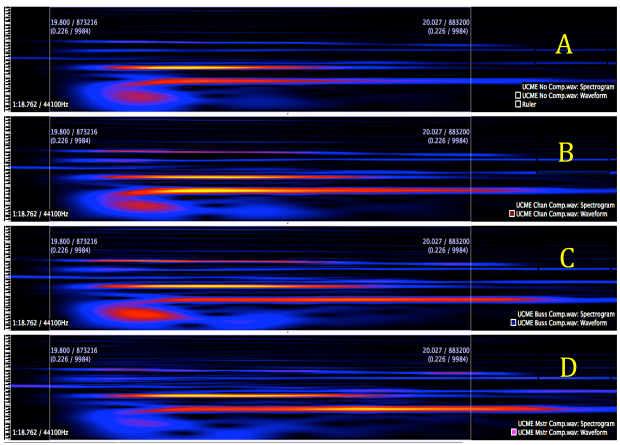


Figure 7 *Spectrograms of the mixed musical instrument signals under test showing the four compression configurations. Each window displays frequency (top HF, bottom LF) vs. time (left to right); Window A – No Compression; Window B – Channel Compression; Window C – Buss Compression; Window D – Master Buss Compression.*

In addition to 'no compression' (Figure 7A), three compression configurations are applied: channel (Figure 7B), sub-mix (Figure 7C) and master buss (Figure 7D). One can observe changes in spectral balance between the configurations, indicated decreased seperation between frequency bands. As the number of signals interacting with one another is increased *before* compression, the distinctions between frequency clusters are lost. This is specifically evident when we compare the un-compressed spectrogram to the master buss compression spectrogram. This reduced spectral clarity is a function of the intermodulation distortion, demonstrated in the Matlab part of the experiment, however here being realised in actual musical signals.

Interestingly, in Figure 8 the peak amplitude of the channel compression configuration compared to the uncompressed version is raised by 1 dB. The sub-mix version, Figure 9, peaks at the same level as the uncompressed version, however, this is a very small peak of the guitar, the overall peak signal level of the guitar has a 1 dB reduction from the uncompressed version.

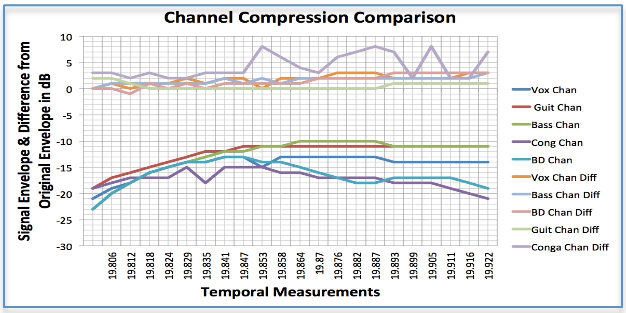


***Figure 8** Measurements of 1 s envelopes of instruments compressed at the channel stage, shown in the lower plots. The plots at the top of the window demonstrate the difference of the compressed signal from the original signal.*

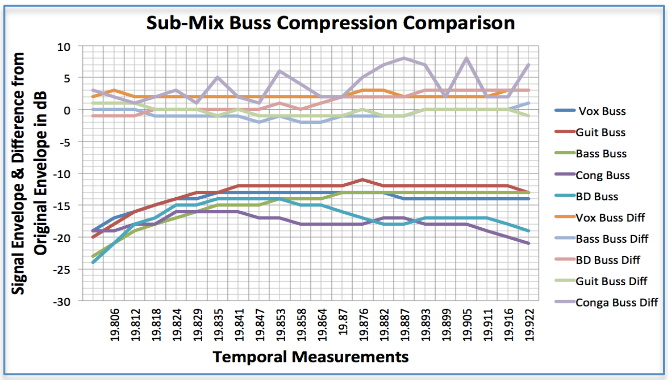


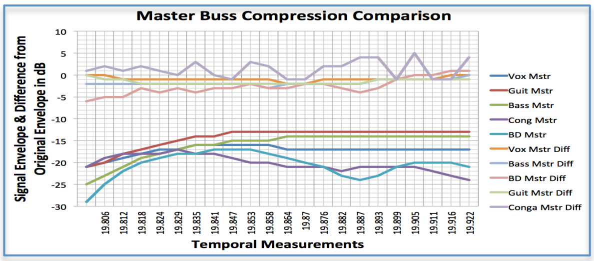
***Figure 9** Measurements of 1 second envelopes of instruments compressed at the sub-mix buss stage, shown in the lower plots. The plots at the top of the window demonstrate the difference of the compressed signal from the original signal.*

It is clear that each compressed configuration does indeed reduce the dynamic variances of the mixed signals. Measuring the difference between the lowest amplitude and the highest amplitude, the un-compressed configuration has a dynamic difference of 27 dB. The sub-mix buss is the least compressed with a dynamic difference of 24 dB, followed by the master buss, Figure 10, with a difference of 16 dB. The configuration that has the least difference is the channel compression, having a signal difference of 13 dB. The channel configuration seems to maintain the closest resemblance to the original un-compressed mix, evident in the overall retention of the original signal envelopes, as well as exhibiting the lowest dynamic difference of the three configurations. It would appear that the channel compression configuration compressed the signals the most, whilst reducing apparent spectral distortions, as indicated by the more defined separations between frequency bands in Figure 7B as compared to Figure 7C and D.

The following observations disregard more subtle (lower amplitude) spectral content, hence timbre, but instead are intended to focus on the most obvious ‘perceived volume’, and indeed evaluative remarks are only in this regard.

The guitar envelope is virtually unchanged by channel compression, apart from the onset and release period. Being that the intended function of compression is to reduce the dynamic variation, it is expected that each of the instruments, apart from the bass drum, have increased amplitude of the onset and release portions of the signal. However, the release period of the envelope is particularly important in relaying the reverberation information of the room and instruments. All of the compressed configurations raise the ‘reverberation envelope’; this is most evident with the bass drum, posibly affect the depth or ‘punch’ of the mix.

***Figure 10** Measurements of 1 s envelopes of instruments compressed at the master buss stage, shown in the lower plots. The plots at the top of the window demonstrate the difference of the compressed signal from the original signal.*



These experiments seem to indicate that the conga envelope is most affected by heavy compression, though it seems channel compression performs best of the three configurations. Distortion of the envelope undoubtedly alters the sound of the drum. The vocals seem to perform equally with sub-mix buss and master buss compression, though this may be misleading since the vocal signal did not share its ‘sub-mix’ buss with another instrument, instead being treated as its own ‘stem’. The channel configuration seems to alter the envelope of the vocals the most, which is an unexpected result, requiring further analysis to fully understand.

The bass also performs best in the sub-mix as its envelope resembles the original most in this configuration. The master buss compression configuration appears to alter the envelope as well as reduce its amplitude the most. This seems to be an example of the lower frequencies being affected by compression as demonstrated in the Matlab part of these experiments. It is commonly understood that the higher energy content of low frequency components drives compressors harder, though this requires further exploration to verify.

After the guitar, it is interesting to see that the bass drum has the second highest amplitude in the attack period of the non-compressed signal set. It is generally accepted that the onset phase of a percussive instrument is most important as it allows the instrument to cut through the mix and establish the ‘beat’. It is therefore significant that the bass drum has such a significant amplitude reduction in the master buss compression signal set, becoming the signal with the lowest amplitude. This appears analogous to the Matlab experiment, which showed a reduction of the signal with the lowest frequency content intermodulating with other signals, again due to typical energies in such bands.

6. Conclusions

It must be strongly emphasised that this research does not advocate using compressors in such an extreme way, but when we can quantify the effects of heavy compression, it is clear that there will also be more subtle ramifications in everyday use, and we can then highlight the importance of responsible practice. Use of dynamic-range compression to increase the overall loudness of music programs is known to reduce dynamic variance. However, it has been shown that compression can be the source of rearrangement of dynamic variances related to the rhythmic structure of musical signals. These rhythmic structures are often the defining aspect of a piece of music, so alteration is tantamount to changing the feeling and emotion of the track, quite possibly to detriment — surely not a desired artefact when actually seeking to increase the perceived ‘impact’ of the piece.

As the Matlab experiments have demonstrated, intermodulation distortion has the effect of filling in the frequency bands around the original frequencies, manifesting itself as noise. The more a signal is compressed, the higher the amplitude of the noise; this may not be so noticeable in the lower frequencies since our ears are less sensitive in this band, but when noise is created at frequencies to which our ears are more sensitive, it becomes more audible.

These experiments have also implied that temporal clarity is also affected by these distortions. Compression has the effect of increasing the amplitude of low-level sounds such as reverberation tails, delay effects, and even the natural release period of instrument envelopes. The increase in the amplitude of these low level sounds effectively seems to reduce the depth and perhaps even width of the stereo field.

This research has shown that decreasing the number of signals interacting with each other whilst applying an equivalent amount of compression can reduce the intermodulation distortions and thereby improve clarity. Also, reduction of the amount of compression applied to a signal reduces the creation of intermodulation artefacts.

Although this research has gone some way to understanding the effects of compression on musical signals, there are still a number of unanswered questions. Further research will use Matlab to automate the process of analysing musical signals and further relating the simple signal structures to musical structures.

8. References

1. Serrà , J., Corral, Ã., Boguñá, M., Haro, M. & Arcos, J.L. 2012, "Measuring the evolution of contemporary western popular music", Scientific reports, vol. 2.
2. Roads, C. 1979, "A tutorial on non-linear distortion or waveshaping synthesis", *Computer Music Journal,* vol. 3, no. 2, pp. 29-34.
3. Dutilleux, P., Dempwolf, K., Holters, M. & ZÃ¶lzer, U. 2011, "DAFX: Digital Audio Effects; Nonlinear Processing ", , pp. 101 <last_page> 138.
4. Campbell, W. 2012, "Dynamic Range Compression Investigations", *CoDE: Cultures of the Digital Economy 1st Annual Conference* Anglia Ruskin University, Cambridge, UK, 27/3/12 - 28/3/12.
5. Izhaki, R. 2008, "Mixing audio: concepts, practices and tools" in , 1st edn, Focal Press, , pp. 272-339.
6. Katz, B. 2007, *Mastering Audio, the Art and Science,* Second edn, Focal Press, Canada
7. Toulson, E.R. & Campbell, W. 2009, "The life and death of dynamic range: who decides how loud?", The Art of Record Production Conference, Cardiff, Wales, .
8. Campbell, W., Toulson, R. & Paterson, J. 2010, "The effect of dynamic range compression on the psychoacoustic quality and loudness of commercial music", INTER-NOISE and NOISE-CON Congress and Conference Proceedings Institute of Noise Control Engineering, pp. 1087.
9. Ward, J. & Campbell, W. 2010, "Loss of Our Musical Heritage; The Rise of the Digital Re-Master", The Art of Record Production Conference 2010 Leeds Metropolitan University, Leeds UK, 3/12/2010 - 5/12/2010.
10. Stone, M.A., Moore, B.C.J., Fullgrabe, C. & Hinton, A.C. 2009 July/August, "Multichannel Fast-Acting Dynamic Range Compression Hinders Performance By Young Normal-Hearing Listeners In A Two Two-Talker Separation Task", Journal of the Audio Engineering Society, vol. 57, no. 7/8, pp. 532-546.
11. Cannam, C., Landone, C. & Sandler, M. 2010, "Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files", Proceedings of the international conference on MultimediaACM, , pp. 1467.
12. Avid | Dynamics III . Available: http://www.avid.com/US/products/Dynamics-III [2013].