Evaluating harmonic and intermodulation distortion of mixed signals processed with dynamic range compression

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. 224-246.

Evaluating harmonic and intermodulation distortion of mixed signals processed with dynamic range compression

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

[1] CoDE Research Institute, Anglia Ruskin University, UK.
rob.toulson@anglia.ac.uk

[2] Faculty of Science and Technology, Anglia Ruskin University, UK.
william.campbell@anglia.ac.uk

[3] London College of Music, University of West London, UK.
justin.paterson@uwl.ac.uk

1. Abstract

Dynamic range compression of simple signals results in harmonic nonlinear distortion. However, for summed signals, predominantly inharmonic intermodulation distortion (IMD) is generated. This research compares the methods of compressing signals prior to and after summation in order to identify approaches to reduce the level of IMD. Results show that lower IMD values are achieved by applying compression prior to summation.

2. Introduction

This research evaluates the audio signal distortion artefacts that are generated during a dynamic range compression process. Distortion for simple audio systems can be described most generally as the difference between the input and output audio signal [1], however it is often necessary to discuss distortion as the *unwanted* differences between the input and output signal. For example, an audio processor may be designed to create sub-harmonic components for enhancing the perception of bass frequencies in an audio signal, but this process might also introduce an unwanted phase shift at higher order frequencies. In this particular example the phase shift can be considered as distortion, although the generation of sub-harmonic frequency components was indeed the intended function of the processor, so this is not generally regarded as distortion. Distortion artefacts are researched and understood in many areas of audio engineering and electronic component design, however this article

specifically evaluates distortion artefacts applicable to the music production process of dynamic range compression, and so discusses common audio engineering terms within a creative music application. This article therefore somewhat bridges the gap between the science and the art of music production, bringing a unique multidisciplinary approach to such research and hence yielding a framework for future knowledge generation and contextualised understanding.

3. Background

3.1. Harmonic and in-harmonic distortion

Moore et al. [2] categorise the two types of audio signal distortion as a) *linear distortion*, which refers to changes in signal amplitudes and the relative phase of frequency components in the original signal and b) *nonlinear distortion*, which refers to the introduction of new frequency components that were not present in the original signal. Moore et al.'s research extends further to define and model the perceptual effects of linear and nonlinear distortion, deducing that linear distortion is perceived generally as changes in timbre, tonality or 'coloration', whereas nonlinear distortion results in listeners describing the output signal with levels of 'harshness' or 'roughness' [3] [4] [5].

If an audio processor's aim is to be a sonically invisible part of an audio signal chain, then the amount of unwanted harmonic distortion can be measured and calculated as *total harmonic distortion* (THD), where harmonic distortion frequencies are quantified at integer multiples of the fundamental audio frequencies. THD is usually calculated from the output signal's frequency spectrum as a percentage based on the ratio of the power sum of all the harmonic components to the power sum of all the harmonics plus the fundamental, as shown in Equation 1 [6] [7].

$$\%THD = 100 * \frac{\sqrt{H_2^2 + H_3^2 + ... + H_n^2}}{\sqrt{H_1^2 + H_2^2 + H_3^2 + ... + H_n^2}}$$

(1)

Where H_1 is the fundamental frequency spectrum power and H_n is the harmonic power of the nth harmonic.

When evaluating a single sinusoid test signal, spectral powers which are not identified as fundamental or harmonic are classified as noise. The noise can also be quantified as a percentage of the fundamental frequency power (N), so allowing the value of THD+N to be

calculated. However, when evaluating more complex signals (in this context, signals with more than one frequency component), it is not sufficient to label all inharmonic content simply as 'noise'.

Some audio processes result in strong distortion components generated at inharmonic frequencies. Although inharmonicity is an essential part of many familiar musical timbres, Katz [8] describes unwanted inharmonic artefacts as sounding 'nasty' and 'severe'. They are also potentially fatiguing to a listener, predominantly because they are not 'in tune' with the fundamental frequencies in the manner of harmonic distortions [7] [9]. If an audio processor generates harmonic distortion when tested with a single sine wave, then it is generally likely to generate inharmonic distortion when tested with multiple sine waves; this is also referred to as *intermodulation distortion* (IMD). When two mixed sine waves are used for testing, the IMD components are seen at sum and difference frequencies of the two test signals [10]. Therefore, if a test signal is made up of two sine waves f_1 and f_2 that are summed together *prior* to dynamic range compression, distortions are expected at integer multiples of each of the two frequencies (harmonic), but also at sum and difference frequencies of f_1 and f_2 and at integer multiples of those sum and difference frequencies [11]. Where IMD is to be quantified, it is possible to use the same form of Equation 1, though in this case using the powers of inharmonic components for evaluation as follows:

$$\% IMD = 100 * \frac{\sqrt{I_1^2 + I_2^2 + ... + I_n^2}}{\sqrt{H_1^2 + H_2^2 + I_1^2 + I_2^2 + ... + I_n^2}}$$

(2)

Where H_1 and H_2 are the fundamental test frequencies' spectral powers and I_n is the nth inharmonic distortion component power found in the processed signal. Note that the analysis of IMD is not limited to only two source signals, so there may be numerous H values. Increasing the number of H values does however make it a considerable challenge to develop tools and models to accurately analyse and detect THD and IMD components of a processed signal.

Evaluation of IMD involves analysis of 'a more reasonable approximation of a real world signal' [12], yet, to date, the effects of inharmonic distortions have not been quantitatively evaluated with respect to dynamic range compression of musical signals.

High quality audio products are designed to have very low distortion values. For example, a typical power amplifier might have a THD rating of 0.01%, whereas a high quality loudspeaker may produce around 0.5% THD. Distortion can also be used as an audio effect, and audio distortion units can regularly achieve 80% THD or more. IMD

values are usually tested for audio products, but are rarely quoted in technical specification data. It could, however, be argued that the IMD data is more valuable to a customer than THD, as it is often regarded to represent a distortion rating more applicable to real audio signals [9].

3.2. Dynamic range compression in music production

Dynamic range compression is an audio process that is used to reduce the range between the highest and lowest amplitudes of a signal, releasing excess headroom that can be utilised through gain. Following such an increase, the compressed signal has a raised RMS (root-mean-square average) power. Compression is applied as attenuation of signals that exceed a set threshold. The compression ratio and threshold can be adjusted, as can the makeup gain and attack and release response times of the processor. Compression is used in many different ways by many audio engineers, e.g. to even the dynamics of a bass guitar track [13], to help a lead vocal 'sit correctly' in a dense audio mix [14] or to increase the loudness of a sub-mix or complete audio track [15]. However, a number of music producers, for example Alan Parsons [16], avoid the use of compression completely owing to a desire to keep a high dynamic range and to minimise the number of unwanted artefacts created.

Dynamic range compression does indeed lead to noticeable distortion artefacts when used aggressively. While setting the compressor parameters, the sound engineer makes a judgement on the trade-off between the desired compression effect and the unwanted distortion artefacts that are generated. It can be seen that as a single sine wave is compressed, harder compression results in the sine wave becoming increasingly more like a square wave (or 'clipped'), as shown in Figure 1. In Figure 1, for demonstration purposes, we refer to 'soft' compression as being a low compressor ratio applied at a low threshold, whereas 'hard' compression is seen as a high ratio applied at a high threshold. The particular compressor algorithm used to gain results for Figure 1 operates on peak sample values (as opposed to a windowed average sample value), and with very fast attack and release times (i.e. faster than the data sampling period).

The discrete Fourier series expansion for a square wave shows that odd harmonic components of the fundamental square wave frequency are evident [17], as given by Equation 3, so the process of compression on a single sine wave is expected to generate odd-harmonic nonlinear distortion, and this result is verified by Figure 1, which shows the generation of odd harmonic distortion components.

$$y = \frac{4}{\pi}\cos(2\pi ft) - \frac{4}{3\pi}\cos(3*2\pi ft) + \frac{4}{5\pi}\cos(5*2\pi ft) - \frac{4}{7\pi}\cos(7*2\pi ft) + ...$$

(3)

Where y is the signal amplitude, f is the fundamental frequency (Hz) and t is time (seconds).
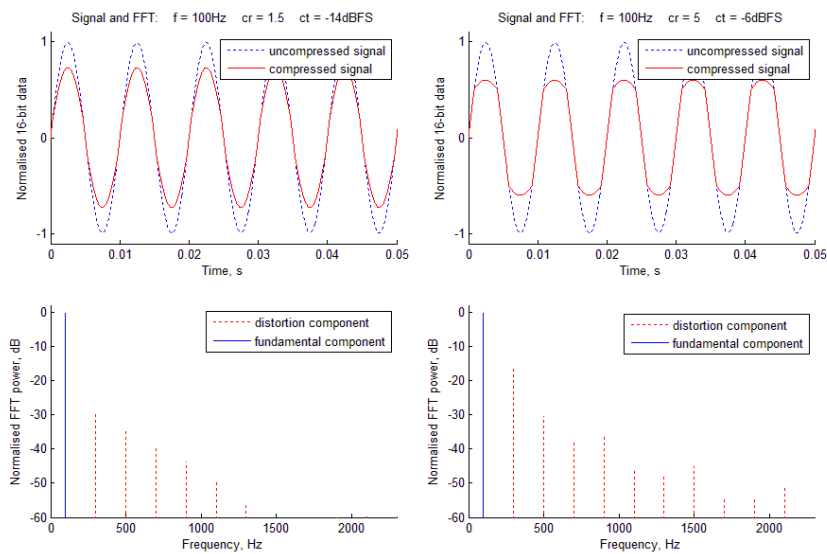

Figure 1. Simple sine wave compressed soft compression (left) and hard compression (right).

Harmonic distortion is often regarded as being 'in tune' with the source audio and therefore does not sound particularly unpleasant at low levels [18]. Indeed, a number of audio *enhancer* effects are available which actively introduce higher-order harmonics to the audio signal in order to accentuate its presence in a mix [19].

3.3. The loudness war

The generation of both types of distortion components is of particular interest given the extensive and ubiquitous use of dynamic range compression in modern commercial music production. Evidence of a commercial *loudness war* has been repeatedly reported showing gradual increases in RMS levels of produced music over a number of years [8] [20] [21] [22]. Record companies often contribute to this

trend with a desire to release louder songs, which are shown to have more immediate impact for the listener. However, there has been a backlash from listeners and music producers recently since such highly compressed audio may cause a subjective phenomenon referred to as *listener fatigue*, as described by Rumsey [23]:

'The constant quest for greater loudness, an obsession with pushing levels to the maximum, and a lack of understanding of the ways in which digital equipment can generate distortion all seem to lead to an increase in listener fatigue.'

The phenomenon of 'listener fatigue' is regularly referred to, however to date, no quantitative study has shown explicitly how and when it is caused; further research into its precipitation and related psychoacoustic factors is clearly required.

Vickers [24] has conducted a thorough review of the loudness war and concludes that:

'While loudness may already be irrelevant to listener preference and commercial success, this idea is contrary to widespread and firmly held assumptions in the (music) industry.'

Well-publicised initiatives by the European Broadcasting Union have been implemented to normalise program loudness for radio and TV broadcast [25]. Here, particular issues include both loudness between stations and also between programs [26], however very little action has been taken to date in a similar respect to commercial music releases. The commercial music production process is indeed a unique one in that dynamic range compression might be applied at many different points during production and distribution. Firstly at the recording and mixing of individual tracks, then perhaps with in-line and/or parallel bus compression [8]. The final audio is often compressed during mastering [27] and then a further time during radio broadcast or playback. This process therefore involves a number of signal summation and compression processes that have not previously been evaluated with respect to the generated distortion characteristics. Indeed, very little *quantitative* data has been generated for formally evaluating the impact of the loudness war in terms of signal artefacts, or with regard to forming an objective definition of listener fatigue. This article therefore evaluates the order in which summation and dynamic range compression are applied in music production processes with specific reference to the types of distortion generated, and offers a quantitative perspective of these distortions. The psychoacoustic analysis of listener fatigue is not within the scope of this article, however the results obtained will be valuable for future studies on the evaluation of listener fatigue.

4. Research Approach

It is clear that signal processing research can be conducted to further quantify the effects of the loudness war on the audio signal. Stone et al. have embarked on a number of psychoacoustic investigations related to dynamic range compression with respect to listener fatigue, but to date only speech has been used as the test material [28] [29]. Stone et al.'s research shows that dynamic range compression does hinder performance when subjects were completing a speech separation task, indicating that induced distortion artefacts or the onset of listener fatigue was having a detrimental effect on the listener. The research presented herein discusses one particular aspect associated with the loudness war; quantitative analysis of nonlinear distortion induced by dynamic range compression processes. In this article, research is only conducted with simple sine waves as the source input data, but the obtained results provide knowledge to underpin future research focusing more specifically on source data using both real and synthesised music signals.

4.1. Research objectives

The research analysis is separated into three Research Objectives:

(RO1) To verify and quantify the existence of harmonic (THD) and inharmonic (IMD) distortion when compressing mixed signals.

(RO2) To evaluate the generation of THD and IMD with respect to compressor threshold and ratio settings.

(RO3) To evaluate the relative levels of THD and IMD with respect to the signal path position of compression relative to summation.

4.2. Compressor topology

The Matlab compressor design used in this research is shown in Figure 2. Signals exceeding a threshold value (positive or negative) are attenuated by a compression ratio for the proportion of the signal that exceeds the threshold, i.e.

$$attenuation = \frac{1}{cr}$$

(4)

where cr is the quoted compression ratio and is always greater than or equal to 1. The algorithm does not modify data samples that do not

exceed the compressor threshold. In this research, compressor makeup gain is not employed in order to avoid adding energy to the compressed signals; this ensures that distortion components are not boosted unrealistically, though in a real world application makeup gain is often used to utilise the additional signal headroom. Test signals are made up of summed sinusoids of equal amplitude and the resulting test signals are normalised to have a peak amplitude of –0.1 dBFS, where 0 dBFS represents the maximum possible full scale value. More complex and analogue modelled compressor designs are discussed by Giannoulis et al [30], however, for this research the compression algorithm is designed simply to be fast acting based on immediate peak sample values, rather than to utilise RMS (average) threshold response or attack and release time designs, which impart their own filtering characteristics on the processed signal.
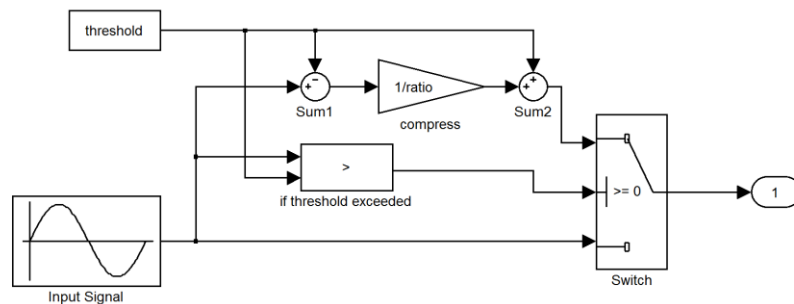

Figure 2. Compressor topology for testing.

The algorithm is run on 10 second duration 16-bit test signals generated at a sampling frequency of 44,100 Hz. Where frequency analysis is performed, the built-in Matlab fast Fourier transform (FFT) function is used to give a frequency resolution of 0.1 Hz. The values of THD and IMD are calculated by implementation of Equations 1 and 2 respectively, and these are used to evaluate the levels of distortion encountered through a compression process. To avoid quantisation and noise readings during calculation of THD and IMD, only spectral peaks above –60 dB and up to a frequency of 10,000 Hz are used for analysis.

5. Research Analysis

5.1. Identifying distortion characteristics of the compression process (RO1)

Dynamic range compression is a nonlinear process, so it is expected that THD and IMD components will be generated when a signal is

manipulated in this way. Where intermodulation distortion is present, the IMD components are expected at sum and difference frequencies of the mixed input test signals.

To verify the relationship between fundamental and distortion components, a number of signals, each made up of two mixed sinusoids are compressed as shown in Figure 3. Research Objective 1 is investigated here using fixed compressor settings, to allow direct comparison of distortion generation as the frequencies of the input signals are varied.
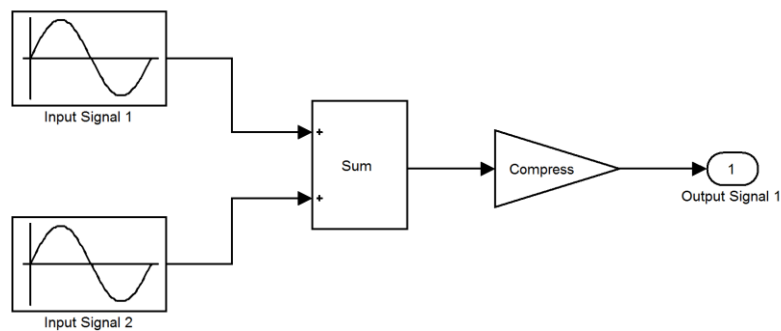

Figure 3. Research Objective 1 evaluated by summation of signals prior to compression.

The compression ratio used is cr = 2 applied at a threshold of ct = −6 dBFS. Frequency f_1 is fixed at 100 Hz and the value of $df = f_2 − f_1$ is adjusted to give a representation of the relationship between the two frequencies and the distortion components resulting from dynamic range compression.

Initially, the evaluation considers distortion components where f_1 and f_2 are harmonically related. A harmonic relationship is defined by the equation $f_2 = kf_1$, where k is a positive integer. Figure 4 shows the test signal and the signal FFT both before and after compression for f_1 = 100 Hz and f_2 = 200 Hz (i.e. k = 2 and df = 100). As expected, distortion components are seen only at harmonic values of the two fundamental frequencies. The time domain signal affect of the compressor shows a considerable re-shaping of the waveform, resulting in THD of 9.44%. It is also verified that IMD = 0% on all occasions where the input frequencies are harmonically related.
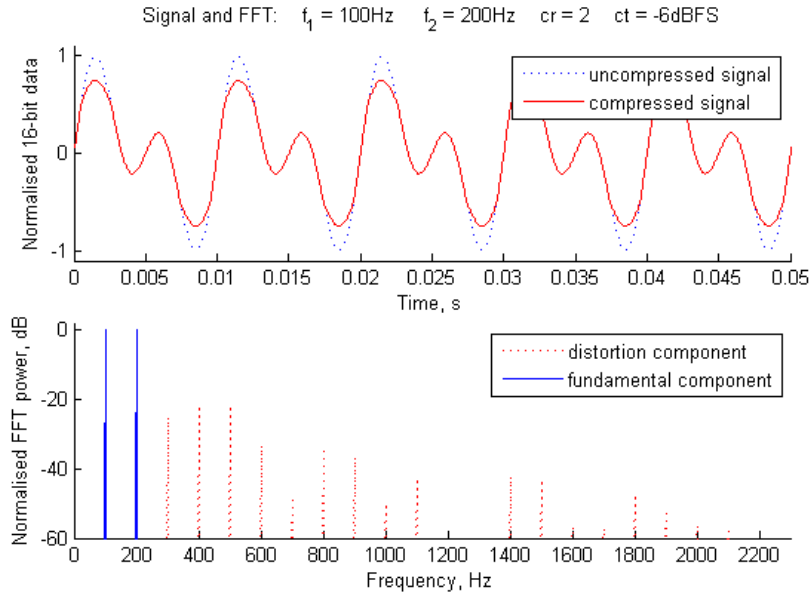
Figure 4. Purely harmonic distortion seen when compressing a signal of harmonically related summed sinusoids (THD = 9.44% IMD = 0%).
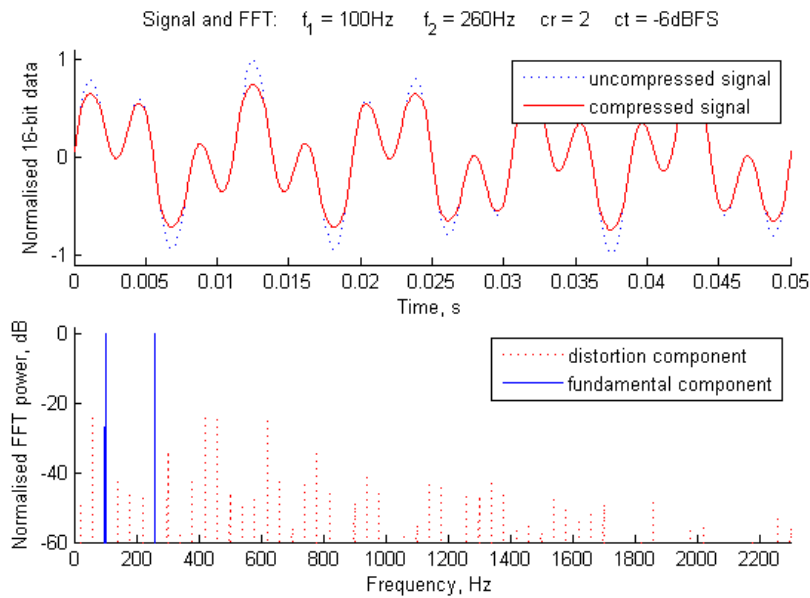

Figure 5. Distortion components for f1 = 100 Hz and f2 = 260 Hz (THD = 2.20% IMD = 9.75%).

The case where f_1 and f_2 are not harmonically related is now considered. For example, Figure 5 shows the test signal and the signal FFT both before and after compression for arbitrarily chosen frequencies f_1 = 100 Hz and f_2 = 260 Hz (i.e. df = 160). Here we can see that the four strongest distortion components are at 60 Hz, 420 Hz, 460 Hz and 620 Hz. There is considerably less THD than in Figure 4 (2.20%), though a considerable reading for IMD (9.75%).

The four most prominent distortion components are analysed for a number of different inharmonic df values and are shown in Figure 6, which highlights four clear linearities with two different gradients.

The equations for the four linearities are extracted from Figure 6 and are given in Table 1, which shows a portion of the calculated distortion component data. This experiment is also verified for a number of different f_1 values.
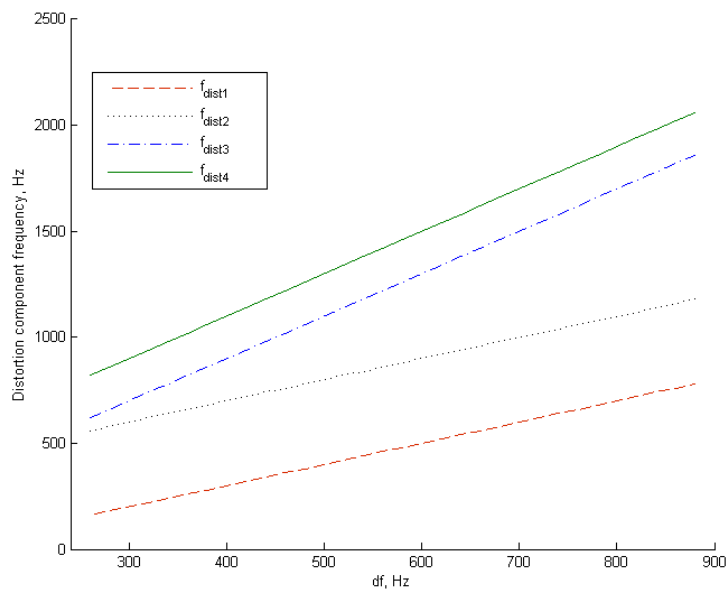

Figure 6. Distortion components for a compressed signal based on two summed sine waves, where df = f_2 – f_1

f_1	f_2	Df	f_{dist1}	f_{dist2}	f_{dist3}	f_{dist4}
100	420	320	220	620	740	940
100	440	340	240	640	780	980
100	460	360	260	660	820	1020
100	480	380	280	680	860	1060
100	520	420	320	720	940	1140
100	540	440	340	740	980	1180
100	560	460	360	760	1020	1220
100	580	480	380	780	1060	1260
100	620	520	420	820	1140	1340
Relationship			$f_2 - 2f_1$	$f_2 + 2f_1$	$2f_2 - f_1$	$2f_2 + f_1$

Table 1. Data showing the most prominent distortion components and their relationship to f_1 and f_2 (all frequencies in Hz).

The tabulated data verifies that IMD is generated when mixed signals are processed with dynamic range compression and also verifies that the IMD components are found at the expected sum and difference frequencies. There is an important outcome here for consideration in a music production process, given that real musical signals are made up of complex combinations of harmonic and inharmonic components. It is shown that applying dynamic range compression to a mixed signal of inharmonic components results in the greatest distortion powers being at inharmonic intervals, which as discussed earlier can sound 'nasty' and 'severe'. Indeed, the design of digital-to-analogue convertors for audio is particularly concerned with reducing IMD artefacts encountered in the conversion process. Furthermore, it has been seen that the distortion levels encountered for a moderate compression process are relatively large for complex signals (i.e. around 10%).

5.2. Evaluating THD and IMD with respect to compressor settings (RO2)

Research Objective 2 involves the evaluation of distortion when manipulating compressor threshold and ratio settings in compression of a signal made up of two summed sine waves. This experiment also uses the signal flow topology described by Figure 3.

Fundamental frequencies, which are not harmoniously related, are again chosen as 100 Hz and 260 Hz, although this experiment has been repeated with different fundamental frequencies (f_1 and f_2) showing similar results.

Figure 7 shows the results for varying the compression ratio whilst holding a fixed compressor threshold of –6 dBFS. Figure 7 clearly shows that no distortion is present at a compression ratio of 1 (i.e. no compression applied), but, as expected, the distortion levels increase as the compression ratio increases. The distortion levels asymptote for both harmonic (THD) and inharmonic (IMD) distortion, and IMD is seen to be approximately 4 times greater than THD at all times. The asymptote is owing to the fact that the compressed signal becomes hard-clipped at higher ratios, but the effect of increasing the ratio further has a diminishing effect on the level of clipping and hence the amount of distortion.

The results for evaluating distortion generation against compressor threshold (with fixed ratio cr = 2) are shown in Figure 8 and bring a less obvious result. Here it can be seen that as the threshold is lowered from 0 dBFS, unsurprisingly the distortion levels for THD and IMD increase. However, the IMD has a peak distortion level at approximately ct = –10 dBFS and from here onwards the distortion levels decrease as the threshold is lowered. The reason for this is that applying compression with a fixed compression ratio acts predominantly as a simple gain effect at very low thresholds, because it attenuates the majority of the signal waveform equally, so distortion is relatively low. At high thresholds the compressor acts more like a clipping process, so distortion levels increase as the threshold is reduced and more of the signal is affected. There is, however, a point where a trade-off between the two effects is encountered, resulting in the peak value observed in Figure 8.
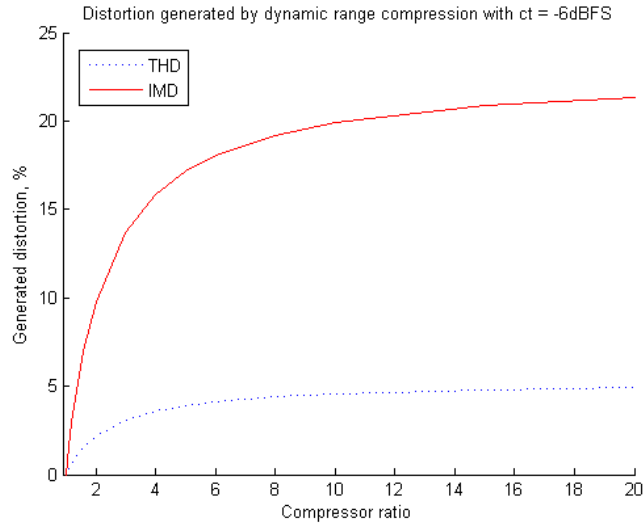
Figure 7. Distortion induced by compressing a mixed signal with f1 = 100 and f2 = 260 with variable compression ratios. Compression threshold fixed at ct = –6 dBFS.
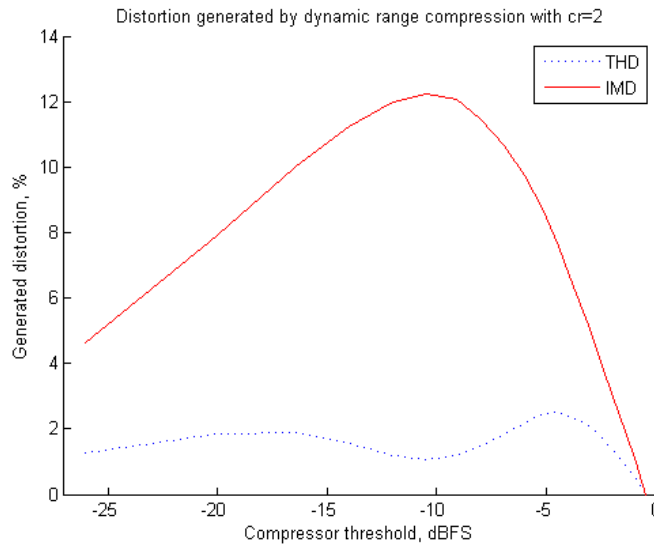

Figure 8. Distortion by compressing a mixed signal with f1 = 100 and f2 = 260 for variable compression threshold. Compression ratio fixed at cr = 2.

This is an interesting observation that could provide rationale for a phenomenon already implicitly understood by music producers and mix engineers. The result indicates that to dynamically compress an audio signal whilst avoiding high levels of distortion (<4%) the compressor might be used in one of two ways:

1. As a high ratio compression process with a high threshold
2. As a low ratio compression process with a low threshold

Clearly a high ratio with low threshold results in very heavy compression and substantial distortion artefacts, whereas a low ratio with high threshold results in only a minimal compression effect.

The two suggestions of best practice for implementing dynamic range compression with the minimum distortion artefacts are not thoroughly investigated here, however the initial observations clearly justify further detailed research in this specific area.

It is also interesting to note that the THD graph in Figure 8 does not follow exactly the same curvature as the IMD graph. This cannot be explained here without further research, which as mentioned in Section 5, will be the subject of further investigation.

5.3. Evaluating sinusoidal distortion components with respect to the position of the compression in the processing chain – sine wave input (RO3)

Research Objective 3 evaluates the levels of THD and IMD with respect to the position of the compression operation within the processing chain. The two mix scenarios described by Figure 9 are considered; the level of distortion is measured for signals that are a) compressed prior to summation (sometimes referred to as *source compression*) and b) summed before being compressed (known often as *bus compression*), and in addition are migrated to a context more closely related to music. This is a novel experiment building on previous research by Campbell et al [22], Stone et al [29] and Giannoulis et al [30].

In this experiment, the number of input signals are gradually increased to show how the distortion artefacts manifest as the source material becomes more complex. The generation of THD and IMD are evaluated for a fixed compression ratio of cr = 2, whilst threshold is varied. The compression ratio of 2 is chosen to allow a relatively mild compressor setting that still generates sufficient distortion components for analysis. Initially we look at three input signals as shown in Figure

9. All input signal are chosen to be inharmonic, because, as shown earlier, when the input signals are harmonically related, no IMD is generated. It must be noted also that real audio signals are not entirely harmonically related either, so this approach best represents a simple model of a real audio mixing scenario. The three input signals chosen are 100 Hz, 170 Hz and 260 Hz.
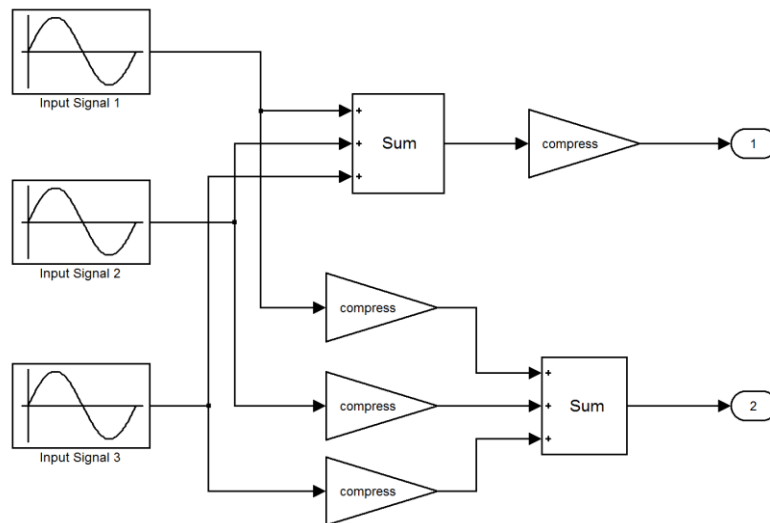

Figure 9. Experimental design to evaluate the impact of the position of compression and summation in the processing chain (example shows three source signals).

Distortion levels are analysed against compressor threshold for both source-compression and bus-compression techniques. The results are shown in Figure 10. Figure 10a shows a similar result to that of inharmonic signals summed in Figure 8, however, the source-compression experiment (Figure 10b) shows very different distortion characteristics to Figure 10a.

When source-compression is used on the three input signals, no inharmonic distortion (IMD) is generated at all; this is expected given that the compression is applied to single sinusoids prior to summation. Harmonic distortion (THD) is generated in a similar profile to the IMD of Figure 10a, but with a lower level of distortion.

The experiment is extended to incorporate a total of eight input sine waves. These are at the inharmonic frequencies 100 Hz, 170 Hz, 260 Hz, 310 Hz, 350 Hz, 430 Hz and 470 Hz, which have been arbitrarily

chosen for this experiment. The results for the eight input test are shown in Figure 11.
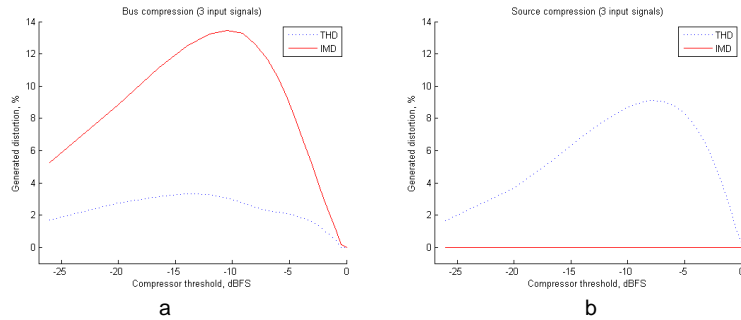

Figure 10. Comparison of THD and IMD distortion levels for a) bus-compression and b) source-compression techniques (three input signals).
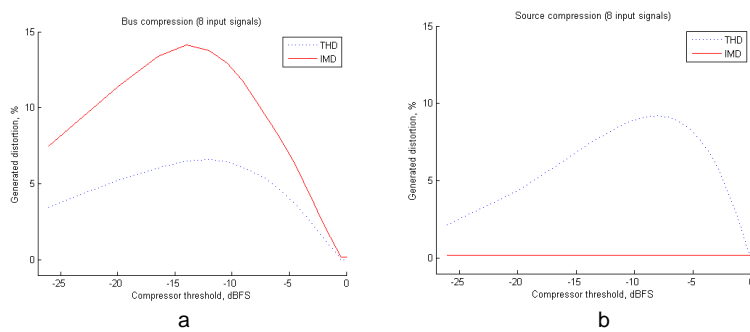

Figure 11. Comparison of THD and IMD distortion levels for a) bus-compression and b) source-compression techniques (eight input signals).

Figures 10 and 11 give an interesting result which may be of value in informing the music production process. Although real audio signals are more complex than simple sine waves, these results indicate that for IMD levels to be kept to a minimum, dynamic range compression should ideally be implemented at an early stage in the summation process.

The total distortion levels for each process can also be evaluated by summing THD and IMD for the two summation and compression arrangements. These results are shown in Figure 12.
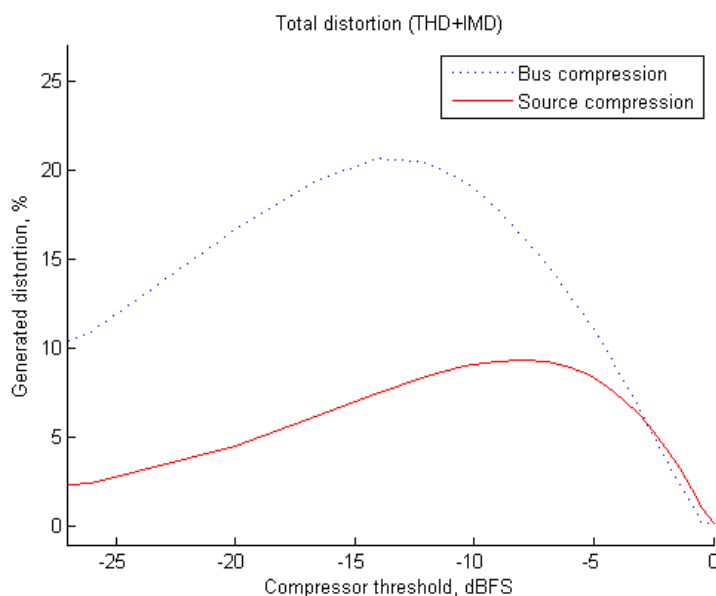
Figure 12. Comparison of total distortion levels for buss-compression and source-compression techniques (fixed ratio cr = 2).

The results shown in Figure 12 can also inform the music production process. Here it can be seen that if total distortion is to be kept to a minimum, then source-compression is preferable for hard compression requiring low threshold values. If high-threshold compression is required then bus compression might appear to offer a benefit, although it must be noted that although the overall distortion levels are marginally lower (above approximately ct = –3dB), the relative levels of IMD are higher than THD (as shown in Figures 10a and 11a). It may be preferable to keep IMD as low as possible, so source compression may still prove to be subjectively 'cleaner' at all threshold levels.

For completeness, this experimentation has also been repeated with a fixed threshold and variable compression ratio. The results for total distortion are shown in Figure 13.

Figure 13 shows that distortion levels evaluated against compression ratio are always lower for the source compression technique. It should also be noted that in this context, the source compression distortion is entirely harmonic as IMD = 0 for all results. Distortion owing to bus compression is predominantly inharmonic, in similar proportion to that shown in Figure 7 previously. This result further supports the notion that, where possible, source compression is preferable to bus

compression if distortion levels are to be kept low and predominantly harmonic. It must be reiterated however, that this research is specific only to simple sine wave analysis and further research is required to fully evaluate these results with respect to real audio source material.
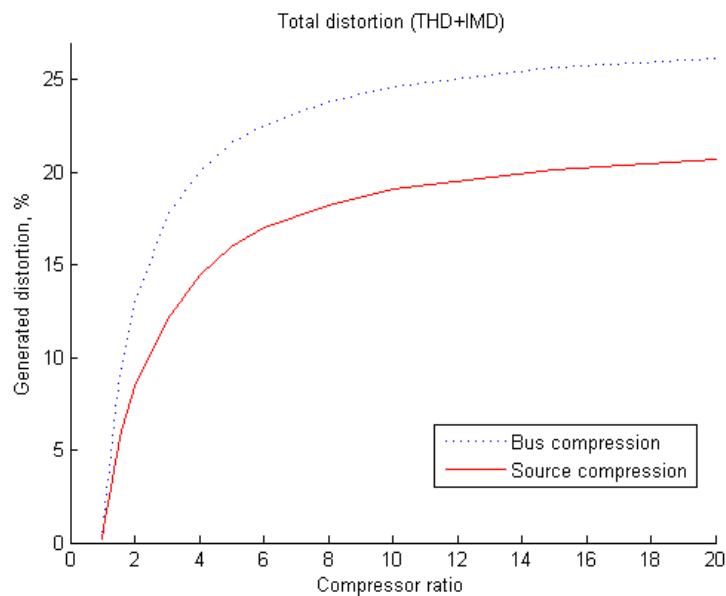

Figure 13. Comparison of total distortion levels for bus-compression and source-compression techniques (fixed threshold ct = –6dBFS).

6. Discussion of current and future research

This research has initially shown that dynamic range compression does, as expected, generate significant levels of nonlinear distortion, measured here as THD and IMD (Research Objective 1). This is an expected result given previous research (discussed by [13], [14] and [16] for example); however, here the implications for music production practice are considered also. Dynamic range compression is a valuable tool for a music production engineer, used in many applications on individual instrument tracks and on mixed audio. The nonlinear distortion generated as part of a dynamic range compression process is usually undesired, and an engineer must weigh up the advantages and disadvantages of a dynamically compressed signal against the distortion artefacts generated – normally a purely intuitive and subjective process.

Research Objective 2 was to quantify the THD and IMD generated for various compressor threshold and ratio setting. This has been investigated particularly for inharmonically related input signals, as harmonically related signals are shown to only generate THD components. The threshold test results showed that for a fixed compression ratio, the level of distortion increases greatly as the threshold is lowered, to a point where the distortion starts to reduce again. This is because a lower threshold affects the entire signal more uniformly, so the process acts more like simple attenuation at very low thresholds. When increasing the compression ratio at a fixed threshold the distortion values tend towards a limit.

This research has also evaluated the placement of a compression process in a simple audio mixing setup (Research Objective 3). When compression is applied after summation the distortion levels are significantly greater and are predominantly inharmonic. This indicates that in order to minimise the amount of inharmonic distortion resulting from dynamic range compression, it is beneficial to compress signals at an early stage in the mixing chain, before summation. This recommendation is contrary to the modern techniques for audio mastering where dynamic range compression and limiting is frequently applied on the final mixed audio in an aggressive fashion. It is suggested that to keep distortion components to a minimum it could be better to compress each individual source audio file prior to mixing, rather than compressing the summed audio after mixing. These results however are only valid for simple sine wave input signals, and further research is recommended evaluating the placement of compression with respect to real and synthesised audio input signals.
A further strand of research could investigate to what extent multi-band compression (typical of the mastering chain) ameliorates such distortion. Further to this, the generation of these distortions might form the quantification of the mysterious 'glue' that is so often sought when compressing an entire mix, and so might not always be undesirable.

At present however, there is a commercial challenge in achieving the source compression mixing strategy, particularly given that it relies on multiple instances of compression processes. Furthermore, audio level metering tools at this time do not sufficiently simplify the process of utilising multiple compression instances. There is therefore an opportunity for new audio processing and analysis tools to be developed to facilitate the source compression process. Furthermore, novel metering tools could be developed to allow mix engineers to analyse their signal dynamic range much more thoroughly at the mixdown stage.

The hypotheses evaluated above, i.e. that the method of source compression for dynamic range control results in lower inharmonic

distortion, than when employing bus compression or master compression methods, is proven for simple sine waves processed through a digital compression algorithm. It is intended also in future to test this hypothesis for more harmonically complex waveforms, real audio source material and in a fully analogue processing and summation setup. Other compressor parameters and indeed modes of operation, including parallel compression techniques, similarly should be investigated. The authors are currently in the process of conducting this extended research.

A further area for enhanced research is towards a formal understanding of 'listener fatigue' with respect to dynamic range compression. Amongst other aspects, the effect of inharmonic nonlinear distortion of audio data can be explored in relation to the onset of listener fatigue. However, this must be conducted in a controlled psychoacoustic test environment, particularly expanding on the previous research studies conducted by Moore and Stone amongst others [2], [3], [4], [5], [25], [26]. Further still, the effect of phase on the generated distortion components has not yet been evaluated with respect to listener fatigue.

This research therefore acts as an initial objective evaluation of compression techniques when applied in an audio production and mixing process. Only nonlinear distortion effects have been evaluated and of course there is a wider discussion on the level and density of distortion components, and indeed the effect of the reduced dynamic range itself on the quality of the listener experience and the onset of listener fatigue. It is felt that there is still a great deal of research and development to be conducted to fully understand the implications of dynamic range compression, listener fatigue and indeed the loudness war. In particular, it is hoped that this research can develop, and generate opportunities for new tools and understanding to improve the objective quality of produced audio and perhaps even the development of effective distortion removal & dynamic range reduction reversal (expansion) tools.

7. Conclusions

Conclusions of this research are summarised as follows:

- Dynamic range compression of mixed signals results in the generation of inharmonic nonlinear distortion components.

- Dynamic range compression applied prior to summation results in fewer harmonic and fewer inharmonic distortion components in the processed signal than if compression is applied after

summation. A recommendation is made for audio and music production that consideration is given for dynamic range compression to be conducted at an early stage in the mixing chain, prior to the summation of source material.

- Opportunities exist for the development of tools to implement source compression and to evaluate dynamic range in a more detailed and holistic manner throughout the audio mixing process, reducing the temptation for heavy compression or lower-threshold limiting at the mix bus or mastering stage of music production. Opportunities also exist for development of effective distortion removal and novel dynamic range expansion tools.

- This research can be enhanced further by considering simple waveforms with internal harmonic structures, real audio material, more extensive evaluation of compressors and processing in the analogue domain.

- The onset of listener fatigue with respect to inharmonic nonlinear distortion as a result of dynamic range compression should be evaluated to obtain an objective understanding of any detectable fatigue and therefore provide an enhanced understanding of the implications of 'the loudness war'.

8. References

[1]   Simpson. A. The Role of the Recording Engineer ('What sounds good, is good?'), Proceedings of The Art of Record Production Conference, Leeds, UK. (2010 December).

[2]   B. C. J. Moore, C.-T. Tan, N. Zacharov, and V.-V. Mattila, Measuring and predicting the perceived quality of music and speech subjected to combined linear and nonlinear distortion, J. Audio Eng. Soc., Vol. 52, No. 12, pp. 1228–1244, (2004 December).

[3]   B. C. J. Moore and C. T. Tan, 'Perceived Naturalness of Spectrally Distorted Speech and Music,' J. Acoust. Soc. Am., Vol. 114, pp. 408–419 (2003).

[4]   C. T. Tan, B. C. J. Moore, and N. Zacharov, 'The Effect of Nonlinear Distortion on the Perceived Quality of Music and Speech Signals,' J. Audio Eng. Soc., vol. 51, pp. 1012–1031 (2003 November).

[5]   B. C. J. Moore and C. T. Tan, 'Development and Validation of a Method for Predicting the Perceived Naturalness of Sounds Subjected to Spectral Distortion,' J. Audio Eng. Soc., vol. 52 (2004 September).

[6]   Temme. S. 'Audio distortion measurements,' Tech. Rep., Brüel & Kjaer, (1992).

[7]   Dutilleux, P., Dempwolf, K., Holters, M and Zolzer, U. Nonlinear Processing in DAFX Digital Audio Effects (Zolzer, U. Editor) 2nd edition, Wiley, Chichester (2011) pp. 101-137.

[8]   Katz. B. Mastering Audio 2nd edition, Focal Press, Oxford (2007), p. 220.

[9]   Bohn, D. Audio Specifications technical note, Rane Corporation, (2000).

[10] C. Thomsen and H. Møller, 'Swept Measurements of Harmonic, Difference Frequency, and Intermodulation Distortion', Brüel & Kjær Application Note, no. I5-098 (1975).

[11] Talbot-Smit, M. Audio Engineer's Reference Book, 2nd edition, Focal Press, Oxford (2001) pp. 3-102.

[12] Temme. S. 'Audio distortion measurements,' Tech. Rep., Brüel & Kjaer, (1992).

[13] Owsinski, B. The Mixing Engineer's Handbook, MixBooks, Auburn Hills (1999) p. 49.

[14] Massey, H. Behind the Glass Volume 2: Top record producers tell how they craft the hits. Backbeat Books, San Francisco (2009) p. 154.

[15] Case, A. U. Sound FX: Unlocking the creative potential of recording studio effects, Focal Press, Oxford, (2007) p. 141.

[16] Massey, H. Behind the Glass Volume 1: Top record producers tell how they craft the hits. Backbeat Books, San Francisco, (2002) p. 116.

[17] Stroud K. A. Further Engineering Mathematics, 3rd edition, Macmillan, London (1996), p. 849.

[18] White, G.D, Louie, G. J. The Audio Dictionary, University of Washington Press, Washington (2005), p. 138.

[19] White P. Creative Recording – part one, effects and processors, 2nd Edition, Sanctuary, London (2003) p. 78.

[20] Fey, F. March for Quality, Studio Magazine, pp. 24-29 (2009 February).

[21] Levine, R. The Death of High Fidelity: In the age of MP3s, sound quality is worse than ever. Rolling Stone Magazine (2007 December).

[22] Campbell, W., Toulson, E. R. & Paterson, J. The effect of dynamic range compression on the psychoacoustic quality and loudness of commercial music, Proceedings of Internoise 2010 Conference, Lisbon (June 2010).

[23] Rumsey, F. Loudness Trumps Everything. J. Audio Eng. Soc., Vol. 54, No. 5, pp.421.423 (2006 May).

[24] Vickers. E. The Loudness War: Do louder, hypercompressed recordings sell better? J. Audio Eng. Soc., Vol. 59, No. 5, pp. 346-351 (2011 May).

[25] EBU R 128-2011. Audio loudness normalisation & permitted maximum level, EBU-UER, Geneva 2011.

[26] Lund, T. Control of Loudness in Digital TV NAB BEC, Las Vegas, United States (2006 April).

[27] Rumsey, F. Mastering in an Ever-Expanding Universe. J. Audio Eng. Soc., Vol. 58, No.1/2, pp. 65-71 (2010 Jan/Feb).

[28] Stone M. A. and Moore B. C. J. Quantifying the effects of fast-acting compression on the envelope of speech. J. Acoust. Soc. Am. 121:1654-1664. (2007).

[29] Stone, M. A., Moore, B. C. J., Füllgrabe, C. & Hinton, A. C. Multichannel Fast-Acting Dynamic Range Compression Hinders Performance by Young, Normal-Hearing Listeners in a Two-Talker Separation Task, J. Audio Eng. Soc. Vol. 57 No. 7/8 pp. 532-546 (2009 July).

[30] Giannoulis, D., Massberg, M. and Reiss, J. D., Digital Dynamic Range Compressor Design — A Tutorial and Analysis, J. Audio Eng. Soc., Vol. 60, No. 6, pp. 399-408, (2012 June).