Hyper-compression in Music Production: The Loudness Normalisation Revolution and Implications for Music Streaming Delivery Platforms.

2016 Conference of ANZCA
University of Newcastle, Australia, July.

Robert W. Taylor
School of Design, Communication and IT, Faculty of Science and IT.
The University of Newcastle
robert.w.taylor@uon.edu.au

Abstract

Music streaming services are creating a new space in the digital world for audiences to connect with artists and acquire music on demand. Innovations such as these, however, present a range of uncertainties insomuch as it is never fully known at the point of formation how the innovation will render antecedent models obsolete and generally interact within the system it is be introduced into. Apart from the fact that music streaming appears destined to be disruptive to the traditional method of physical sales, there have been some other unintended consequences that could not have been predicted. Music streaming in conjunction with a relatively new concomitant technology, *loudness normalisation* (that mediates the perceived loudness of audio content), has been heralded by some within the audio community as a mechanism that will solve a problem that has been endemic in music production practice for two decades; the over use of hyper-compression, dubbed the 'Loudness Wars'. Through the lens of Rogers' 'diffusion of innovations' approach, a theoretical framework is presented to qualify this claim by examining these interrelated technologies, streaming and loudness normalisation, and discuss the possible changes that might occur within the social system of music production and dissemination.

Keywords: Music Streaming, Loudness Normalisation, Hyper-compression, Loudness

Introduction

Loudness normalisation is a relatively new technology designed to address a long standing problem concerning loudness discrepancies between program content in television broadcasting. In simple terms, loudness normalisation allows all program content to be

perceived at the same loudness, negating unwanted discrepancies that may be potentially disruptive to the flow of the overall broadcast (for example, loud advertisements). The technology has also found to be useful for application with various other media platforms. Of particular concern to this study is how the technology has migrated to internet streaming platforms to address similar issues in loudness levels, and interestingly, how it has also been recently argued by some as an unintended but potential solution to another long standing concern in the sphere of music production; the contentious use of hyper-compression dubbed the 'Loudness War" or "Loudness Race" (Katz 2013, pp. 241-256).

When technologies are introduced into systems of cultural production there , as Rogers argues, an 'uncertainty' that surrounds the diffusion of an innovation (2003, p. xx). New technologies can either compliment or become disruptive to existing technologies and practices. It takes time for the success of an innovation to be evaluated, which is 'gradually worked out through a process of social construction' (Ibid). Importantly, Rogers also points out that the implementation of a new innovation can present 'new means of solving problems' that may arise in ways that could not have been predicted (Ibid).

With this in mind, although loudness normalisation has been viewed as a successful solution to concerns specifically related to television broadcast since its implementation in 2011, claims that it may be also effective in changing such a deeply ingrained practice such as the overuse of hyper-compression in music production cannot as yet be confirmed and remains nothing more than a logical assumption. The investigation outlined in this paper examines the possibility of loudness normalisation effectively changing what Rogers describes as the 'structure and function of a social system', evoking what would represent a significant 'social change' (2003, p. 6). To accomplish this task, it is argued that an environment within the

music industry's landscape that abrogates the motivation for audio practitioners to use hyper-compression to make loud masters, must be first in place. It has been suggested that music streaming delivery platforms and personal media players that incorporate loudness normalisation can provide such an environment by removing loudness bias (a predilection by listeners for louder recordings), however the widespread adoption of these technologies is vital for this to occur. The trajectory of these technologies and their adoption rate are therefore of major concern. Accordingly, the investigation herein covers all aspects of the relevant technologies required to potentially create this environment and establish such a significant social change.

A theoretical framework for this investigation is based upon Rogers' theory of the diffusion of innovations. Within this framework, Rogers identifies four main elements as '(1) an innovation (2) is communicated through certain channels (3) over time (4) among the members of a social system', forming the basic structure of the paper (2003, p. 11). It is important to point out that although this process which Rogers suggests may appear to be linear in the way it is presented, in reality it is most certainly not. Despite that the process undoubtedly begins with a problem and a solution, the process by which the innovation develops is a constantly recursive one of re-evaluation and further development. Innovations must adapt in accordance to the uncertainty of innovations that Rogers suggests and considering that loudness normalisation has applications that were not intended in the original design, further consideration and development of the innovation is required. Particular importance is placed on the social system and those members of the social system as Rogers explains:

> Diffusion is a process in which an innovation is communicated through certain channels over time among the members of a social system. Therefore, this element is discussed first as it makes sense to identify who these members are and the roles they play within the social system before embarking on how the innovations works within this system. It is a special type of communication, in that the messages are concerned with new ideas." (2003, p.5)

The investigation explores the interrelationship of key players, technologies and practices. The relevant terminology along with basic principles in audio theory are explained which underpin the technology to provide sufficient background for understanding. An examination reviewing how these interrelated technologies pose a potential solution to the current audio production practise of hyper-compression then follows.

Background on Related Audio Principles

When considering *loudness normalisation* with respect to *music streaming* and how this may affect the use of *hyper-compression* in music production practices, a basic understanding of two underlying principles common to all three topics is required; *dynamic range compression* and *perceived loudness*. A common audio production process known as Dynamic Range Compression (DRC), or simply "compression", is used to manipulate the dynamic range of a sound recording; the difference between loud and soft sections (*figure 1*).

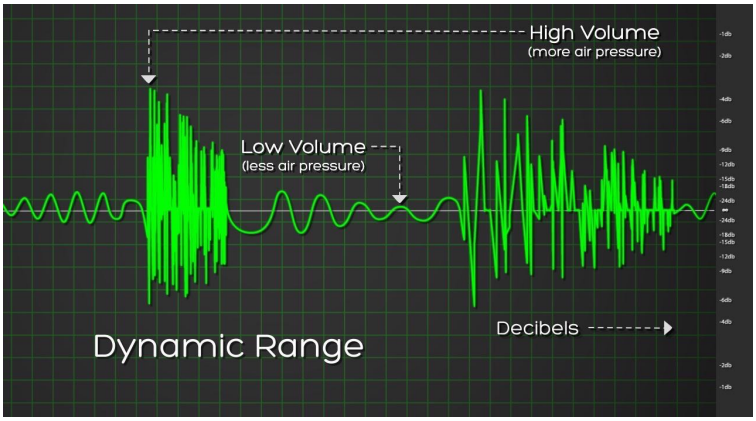

Figure 1: The dynamic range of an audio signal is the difference between the loudest (high level) and the softest points (low level) of a sound as illustrated by the diagram above. Dynamic range control (DRC) or 'compression' can manipulate this range by making the difference between low and high levelsmaller. Image by Michelle Lanier.

This process has myriad applications in audio production, however, specific to this study, it can be used to reduce the dynamic range of an entire music recording. Put simply, compression can reduce loud sections of music and increase soft sections, resulting in all sections of the music having the same loudness level. Hyper-compression is an extreme variant of DRC that can significantly reduce dynamic range and can also increase the overall level of the recording to the absolute highest limit of the media format (*figure 2*). In essence, it makes the music recording extremely 'loud' on playback, especially in comparison to recordings that have not been processed in this way.

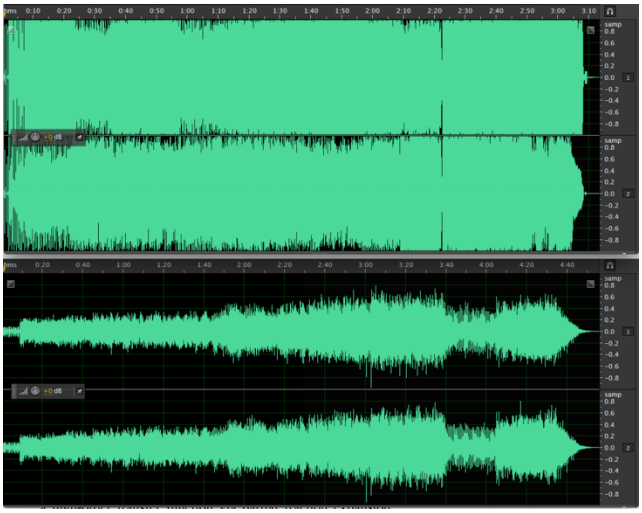

Figure 2: Audio waveforms that depict a hyper-compressed audio signal (top) and non-hyper-compressed signal (bottom). Hyper-compression increases the overall level of the recording to the absolute highest limit of the media format making this recording extremely 'loud' on playback in comparison to recordings that have not been processed in this way. In doing so it reduces the dynamic range the difference between high and low levels.

Perceived loudness is a psychophysical sensation of the intensity of a sound stimulus and represents a basic premise within the study of psychoacoustics[1]. Sound, consisting of fluctuations of air pressure, is transformed into cognitive sensations via the human auditory system, and from there we associate meaning. We can manipulate sound recordings so that

[1] Psychoacoustics is the domain of knowledge related to the behaviour of the human hearing mechanism and how we cognitively interpret sound as a sensation.

the loudness we perceive has a specific outcome, and therefore a specific meaning. Within the context of this discussion, the *loudness* of a sound recording (a common term in the audio community) refers to the level of the recorded signal imprinted on a given media that will eventually correspond to a particular perceived loudness for the listener when reproduced through an audio playback system. Therefore, we can use hyper-compression to ensure that the perceived loudness that is the result of the playback of a sound recording will be extremely loud, that is, be perceived as loud or louder than others thus giving it a competitive advantage.

Alternatively, we may want to make all sound recordings at the *same* perceived loudness for other reasons such as not having jarring jumps in level from one song to another or one TV program to another, hence the use of loudness *normalisation*. Loudness normalisation is a method of correlating audio levels to actual human perception of loudness via an algorithm that in simple terms, mimics the human auditory system with regards to loudness summation and perception (Camerer 2012, pp. 1-9). For all intents and purposes, loudness normalisation makes all program content sound the same perceived loudness, negating unwanted discrepancies in loudness that may be potentially disruptive to the flow of the overall broadcast or general playback.

But that leaves one question. Why would we want to manipulate the loudness of a music recording to make it louder? Research into psychoacoustics has shown that we gravitate towards a preference for music that is perceived to be louder due to inherent structures in the human hearing system; the "louder is better" paradigm (Taylor & Martens 2014, p.1). We can seek answers to why this is the case, first by looking to the research of Fletcher and Munson, who indicated that the human auditory system has a non-linear frequency response at lower

sound pressure levels (SPL) (1932, pp. 82-108). As SPL increases, along with level, our hearing becomes more linear and we perceive more low and high frequencies; we hear greater detail in the sound. The music industry in particular, has actively exploited this cognitive predilection for *louder* since the 1950s and is therefore responsible for large increases in the relative level of sound recordings over many decades (Thiele 2005, p.1). Hyper-compression has now become so ingrained in music production practises that almost all contemporary music recordings are processed in this fashion introducing what has been described as a series of undesirable artefacts such as 'removing dynamics and making music sound "squashed", creating musical clutter and reducing depth and texture, robbing music of its excitement and emotional power' (Vickers 2010, p. 5). However, we mustn't forget that the use of technologies that allow these problems to result occur within a dynamic social system of music production and dissemination.

The Social System

Rogers defines the social system as a 'set of interrelated units that are engaged in joint problem solving to accomplish a common goal' and that these "units" may be 'individuals, informal groups, organizations, and/or subsystems' (2003, p. 23). We can expand these "units" to include all those that are involved in the creation, mediation, dissemination and consumption of a given innovation. The social system, as Rogers explains 'constitutes a boundary within which an innovation diffuses' and within this boundary there are structures that define the behaviour of the social system. There are three basic areas that Rogers outlines here, summarised as the effect of norms (pre-existing structures), the influence of "change agents" on innovation decisions and the final decision whether to adopt or reject the innovation (ibid, p. 24). While it is impossible within the scope of this paper to go into exacting detail as to the scope of the social system participating in loudness normalisation

and music streaming as two separate but related innovations, we can identify key elements within these three areas that are directly related to the use of hyper-compression being the focus of this study.

Firstly, Rogers defines "norms" as the 'established behaviour patterns for the members of a social system' and that 'a system's norms can be a barrier to change' (2003, p. 26). We can establish with certainty that the use of hyper-compression is what Rogers defines as a "norm", a pre-existing structure within the social system that those within the cultural production of music feel bound to use to remain competitive within the market place of music dissemination; a scenario that Weymouth describes as an 'archetypal example of the famous *Prisoner's dilemma* - a theoretical two-player game in which the two players may not co-operate, even if it seems to be in their best interest' (2012, p. 3). With both parties armed with the same technology and ability to make loud masters, a situation developed in the social system not unlike a "cold war" with either party unwilling to back down for fear of being uncompetitive; something that Katz explains as a 'vicious circle of loudness envy' (2015). Vickers takes the concept one step further by stating that:

> The two-player Prisoner's Dilemma does not effectively model a situation like the loudness war, in which there may be dozens, hundreds or thousands of players, each of whom have limited interaction with (and limited leverage over) the others. This type of situation is a "social dilemma," in which individually reasonable behavior leads to a result in which everyone is worse off. As Kollock writes, "a group of people facing a social dilemma may completely understand the situation, may appreciate how each of their actions contribute to a disastrous outcome, and still be unable to do anything about it." (2010, p. 15)

Secondly, we can identify the opinion leaders and change agents in the group known as The Music Loudness Alliance or MLA, who have been at the forefront of the debate *against* the use of hyper-compression and collectively rallied against its use in music production for more than a decade. This group consists of some of the most highly regarded audio practitioners/scientists, globally; Bob Katz, Thomas Lund, Kevin Gross, Bob Ludwig, Florian

Camerer, Eelco Grimm and Ian Shepherd (*figure 3*). Members of the MLA (in particular Camerer and Lund) have also been highly instrumental in the development of loudness normalisation and its implementation in broadcast, which will be discussed later in detail.

Lastly, we can identify the entities within the social system whose innovation-decisions are crucial to change within the structure of the social system with regards to hyper-compression; those involved in the cultural production of music and the music consumer. Music streaming incorporating loudness normalisation is largely viewed as a potential catalyst for this change. If loudness bias is to be removed from music reproduction, the need for the use of hyper-compression will also be removed. Therefore, the adoption of music streaming services which are loudness normalised by the global audience is vital in making these services the dominant form of music delivery. If loudness does not present any advantage to the cultural production of music, as it would on streaming platforms that are normalised, then it is assumed that attitudes of audio practitioners will follow suit. There are, however, a complex array of other factors at play, as suggested by Ronan et al. (2014). Despite the introduction of loudness normalisation, all 'factors contributing to a preference for dynamic range compression have not been removed' (Ronan, Sazdov & Ward 2014, p. 1). Four factors are identified that are suggested to exert influence on the listener apart from loudness bias: 'prolonged exposure to hyper-compressed music; preference for musical genre; education and training; and perceptually salient sound quality attributes' (ibid, p. 4). These factors as well as aesthetic concerns will undoubtedly influence the decision making process by those involved in the cultural production of music. The focus, however, within the scope of this paper is to examine how loudness normalisation and music streaming platforms will exert change within the social system and it is assumed that these other influences on innovation need separate investigation.

Innovation

An innovation, as Rogers argues, 'presents an individual or an organisation with a new alternative or alternatives, as well as new means of solving problems' (2003, xx). There are three related problems we can identify areas in which loudness normalisation as a potential problem solver is associated with its ability to regulate the loudness of audio content:

1. Television broadcast. Discrepancies in loudness, specifically between normal program material and advertising commercials; which was responsible for the largest proportion of complaints worldwide in regards to television broadcast and programming (Eggerton 2014).

2. Internet streaming platforms. Discrepancies in loudness between content and in particular discrepancies between music of different eras; music pre-hyper-compression to that of the hyper-compression era.

3. Hyper-compression in music production. If any difference in loudness between music recordings on playback is removed, then logically the use of hyper-compression and the motivation to create loud recordings would also be removed.

Originating from television broadcast, loudness normalisation is a method for regulating the loudness of program content and prompted the development of a new set of guidelines for standardised loudness. Previously, audio levels in broadcast were regulated using *peak normalisation* in which the highest (peak) level of the signal was not to exceed -6dBFS, but had very little to do with the way we perceive loudness. Although hyper-compressed commercials technically adhered to the required peak level prescribed, they would however be perceived as much louder than more highly dynamic program material (*figure 4*). In

response to the growing concerns over complaints from viewers regarding loud commercials, ones that had been hyper-compressed, the International Telecommunication Union (ITU) instigated standardising the way loudness could be measured and hence regulated in broadcasting. The result became the ITU BS-1770 standard of loudness normalisation which provided a capacity to correlate audio signal levels with perceived loudness (ITU 2012). Therefore, program content of any type could be regulated to potentially yield the same perceived loudness by the listener and although the motivation for this innovation was loud advertisements that caused annoyance to the audience, as mentioned above, it was equally useful for audio content in other platforms.

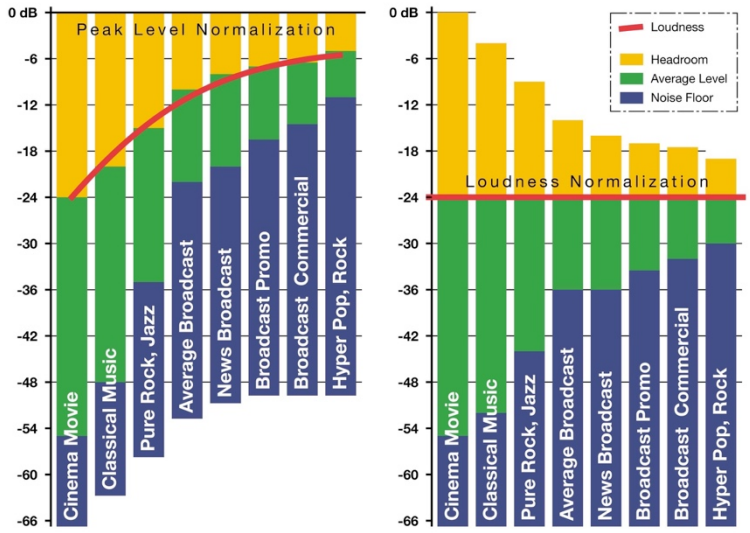

Figure 4: Peak normalisation versus loudness normalisation. The red line indicates perceived loudness. This illustrates that the peak level of an audio signal does not correspond to perceived loudness and the traditional method of using peak meters to regulate the signal level created loudness discrepancies. Loudness normalisation which instead measures the long-term average of the entire audio signal provides a correlation with perceived loudness. Source: Thomas Lund.

Senior sound engineer at the Austrian national public service broadcaster "ORF", Florian Camerer, was given the directive by the European Broadcast Union (EBU) in 2008 to chair a technical committee (known as P-LOUD) to investigate and spearhead the implementation of this new standard. The initial outcome of the committee was the R128 recommendations

published in 2011, of which four supplemental technical documents outlining practical guidelines for its implementation were published in 2011; EBU Tech-3341, 3342, 3343 and 3344 (EBU 2011a-e). These technical documents and recommendations were then referenced by audio manufacturers as guidelines for the development of software and hardware necessary for the implementation of loudness normalisation into production practices (*figure 5*).

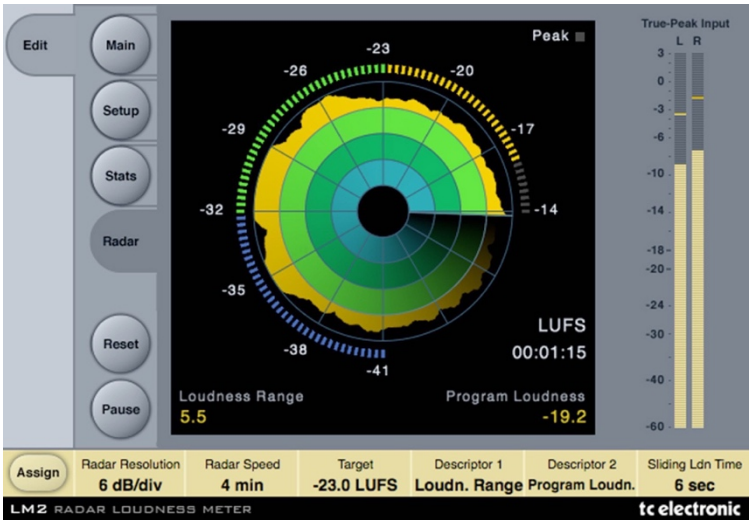

Figure 5: The TC Electronic 'loudness meter'. An example of loudness normalisation technology.

The transformation from peak to loudness normalisation heralded a significant modification in broadcast practises and as Camerer asserts, the use of this innovation was 'even more dramatic than changing from black and white to colour TV' (2015). A new language of audio descriptors was developed, based around measurement units named Loudness Units or LU. Currently, many members of the European Union have become what is now termed "R128 compliant" with their procedures for audio production and broadcast (Camerer, 2015). Countries outside the European Union either developed their own similar version of loudness normalisation practices based on the ITU BS. 1770 standard or adopted the same recommendations as that of R-128. The United States dealt with the problem via legislation; the 'Commercial Advertisement Loudness Mitigation Act', H.R. 1084 (the "CALM Act"); the

regulations enforced by the Federal Communications Commission (FCC) (Powell 2009), and set out in the Advanced Television Systems Committee A/85 recommendations, also based on the ITU BS1770 standard (ATSC A/85, 2013). In contrast to the European model, which was largely voluntary, the United States made it law under which potential fines could apply for non-compliance (Eggerton 2014). In Australia it is also compulsory under the 'OP59' code of practice enforced from January 2013 (Free TV 2012).

Within the ongoing work of the P-LOUD committee, it became apparent that loudness normalisation had far wider implications as Camerer explains:

> …After the ball went rolling and the critical mass had been reached in TV and we saw okay, this is now really going into the right direction, we naturally said the next step is [digital] radio. (2015)

Attention was then directed to digital audio broadcast (DAB), but also to personal media players (iPhone or iPod) and digital streaming services; all of which share a common problem. Music recordings from the decades since the 1950s vary greatly in their overall level and loudness. Modern recordings that have been subjected to hyper-compression can be as much as 20dB louder than recordings from the 1960s to 1980s. This creates a very similar problem to that which television broadcast faced with loud commercials compared to the much lower program material. The change in perceived loudness from song to song can be infuriating to the listener and draws the comparison that the songs of today are hyper-compressed to such an extent that they function *as* "commercials" for the artists that produce them. Loudness normalisation then migrated to most digital platforms that facilitated the reproduction of audio. Several systems were developed for personal media players such as "Replay Gain" and Apple's "Soundcheck", and most streaming platforms such as YouTube, Spotify, Pandora and other DAB radio stations adopted some form of loudness normalisation application that the user could choose to utilise or not. An ensuing problem became evident;

no one standard was applied to any platform and "target levels", one of the key issues in the loudness normalisation process, differed from service to service.

The long-term loudness level of a program, whether it is music, film soundtrack or dialogue, is the actual measure by which a program is estimated. Put simply, the measurement takes into consideration the entire program from beginning to end, as opposed to just the loudest moment; the peak level. The long-term loudness measurement may be significantly less than peak level depending on the dynamic range of the program; described in R128 as the loudness 'centre of gravity' (EBU 2011a). This means that there must be significant headroom available above the centre of gravity to accommodate these peaks to avoid potential overload and distortion. The R128 recommendations stipulate a target level for television broadcast of -23LUFS, which in effect gives 23dB of headroom, enough to accommodate the large dynamic range required for content typical of this media (ibid). Music generally requires less than this: modern hyper-compressed music has a dynamic range of 6 – 10dB; older music up to 16dB; and classical music up to 20dB. The Audio Engineering Society's (AES) technical committee on "Transmission and Broadcast" commissioned a sub-committee to investigate target levels specifically for audio on-line streaming services. The technical document, AES TD1004.1.12-10 - "Recommendations for Loudness of Audio Streaming and Network File Playback," outlines a target level of between -16 and -20LUFS, taking into consideration many factors such as content type and playback device specifications (AES 2015). In conjunction with these recommendations provided by the AES, P-LOUD is also currently drafting another technical supplement to the R128 recommendations that extends to all streaming services including video:

> …all content played sequentially over a network, typically the internet, that's how we define it. Or of course downloads from the net; network file playback directly. So everything that falls into this

> category, Netflix to Beats to Apple, whatever, Spotify, iTunes, everything, Apple Music. (Camerer, 2015)

Another important innovation in this discussion is music streaming (*figure 6*). The music streaming delivery platform has created a new space within the digital realm for audiences to connect with artists and obtain access to their music. It provides an alternative to the traditional ownership of a physical musical artefact, where the audience instead leases music to be listened to on demand. With a similar business model to video streaming companies such as Netflix and Stan, music streaming services provide on-demand access to a network database of music content either for free (with advertising) or as a subscription (without advertising). Music streaming appears destined to be disruptive to the traditional method of physical sales. With over 400 companies operating now in 200 countries, most offering loudness normalised reproduction, music streaming is changing the landscape of music delivery and according to Francis Moore, Chief Executive of the International Federation of Phonographic Industries explains, 'digital services have gone global, enabling the recording industry to reach markets that it could not monetise through physical retailing' (IFPI, 2016, p. 5).

The growth trajectory of music streaming companies on a global scale through its adoption rate by the social system is vitally important to the discussion of hyper-compression. The changes to the 'structure and function of a social system' that Rogers mentions, creating a social change of significance, relies in many ways on how this innovation can be disruptive to traditional models that, it could be argued, had fostered the use of hyper-compression (2003, p. 6). The rationale that loudness normalisation will create an environment whereby sound quality judgements made on music will be made without the bias of loudness, and therefore, negating the need for the use of hyper-compression so as to be competitive in the

market place, hinges on the removal of loudness bias from the listening process altogether. This would require all forms of playback to be loudness normalised, but interestingly, at what level of adoption will the innovation start to have an effect on the social system itself? This is unclear since, as Rogers argues, the uncertainty of innovation and the rate of adoption can be difficult to predict:

> Many technologists believe that advantageous innovations will sell themselves, that the obvious benefits of a new idea will be widely realized by potential adopters, and that the innovation will diffuse rapidly. Seldom is this the case. Most innovations, in fact, diffuse at a disappointingly slow rate, at least in the eyes of the inventors and technologists who create the innovations and promote them to others. (ibid, p. 7)

Although loudness normalised streaming alone is unlikely to solve the issue of hyper-compression, in conjunction with other forms of music delivery and reproduction that incorporate loudness normalisation, it could provide the necessary environment to do so. A key issue here is educating the social system to the potential benefit. Part of that process involves the use of certain types of communication channels.

Communication Channels

Rogers asserts that 'diffusion is a particular type of communication in which the message content that is exchanged is concerned with a new idea' and consequently he suggests the 'essence of the diffusion process is the information exchange through which one individual communicates a new idea to one or several others' (2003, p. 18). In this part of the discussion we specifically address how loudness normalisation is being communicated as a potential solution to the problem of hyper-compression. The fundamental problem with hyper-compression that requires an immediate solution, it could be argued, lies within the recording industry itself and not the consumer as they have little say concerning its use. Moreover, it is the social structure and its actions that produce musical artefacts; the artists, the record companies, the record producers etc. Considering that loudness normalisation is a relatively

new technology, knowledge of its existence and the potential benefit it has is not as yet widely known within the arena of music production, let alone those sectors of the industry that are less technical and more market economy based. It is therefore essential to educate audio practitioners regarding the innovation for any change in the structure of production practices to occur which may then flow on to other sectors. Rogers explains the mechanism of this communication process:

> In its most elementary form, the process involves (1) an innovation, (2) an individual or other unit of adoption that has knowledge of, or has experienced using, the innovation, (3) another individual or other unit that does not yet have knowledge of, or experience with, the innovation, and (4) a communication channel connecting the two units. A communication channel is the means by which messages get from one individual to another (ibid).

Audio Engineering Society (AES) convention seminars have been the main communication channel for the MLA as opinion leaders and change agents to use as a platform for their arguments to members of the audio community. They use this platform to strongly advocate loudness normalisation as a solution to the problem of hyper-compression backed by scientific studies. Whether this very narrow channel of communication can have any effect on the wider music industries and economies is unclear and possibly doubtful. They have, however within this technical community, a voice that is prominent and well respected and may serve as a catalyst for discussion in other sectors. MLA member Bob Katz in particular has proven to be an effective voice in garnering attention to the subject of hyper-compressed recordings since the late 1990s and is widely considered to be one of the leading experts on the subject. In 2013, Katz announced in a press release on his website, 'Digital Domain', that the 'Loudness War is over', citing loudness normalisation and the rise of streaming as a dominant music delivery platform (2013):

> I wanted to be the first to say something about that because at the time I thought that normalisation was happening, that iTunes was going to do this and it was going to change the world […] But it was premature at that time. It was also designed to awaken more people to what normalisation could do. (Katz, 2015)

Katz believes that although the technology is still in its infancy, it's only a matter of time before loudness normalisation permeates the industry and has the desired outcome on production practices (ibid). Bob Ludwig, another high profile member of the MLA, agrees that the recording industry is very slow to take up the innovation of loudness normalisation and anecdotally comments that 'Oh, it's got to be 99.9% of people' that have little to no idea of what loudness normalisation actually is and how it could affect working practices (2015). He also adds, 'well for sure the broadcasters are thinking about it because it's been put into law there', but apart from those that attend AES conferences and read the literature on the topic, it is only slowly making its way into the consciousness of music practitioners.

Apart from a "white paper" titled 'Loudness Normalisation: The Future of Files Based Playback' which outlines an 'integrated solution' for the use of hyper-compression (MLA, 2012), the MLA were also instrumental in devising a set of technical guidelines for music streaming companies mentioned earlier via the AES; The technical document, AES TD1004.1.12-10 (AES 2015). Surprisingly, the MLA have resorted to an on-line petition with activist website change.org in the hope of garnering enough public support to convince music streaming companies to adopt their recommendations as they are not bound to do so[2] (*figure 7*). The AES recommendations are merely that, recommendations but their adoption may only be a matter of time.

[2] https://www.change.org/p/music-streaming-services-bring-peace-to-the-loudness-war

Figure 7: "Bring peace to the loudness war" —the Music Loudness Alliance petition music streaming companies to adopt their recommendation for loudness normalisation.

Time

The 'time dimension (…) involved in diffusion' provides the opportunity to measure the trajectory of an innovation within the social system (2003, p. 20-23). There are two fundamental processes that Rogers outlines these being the *innovation-decision process* in which a decision is made to adopt or reject the innovation, and the *rate of adoption* 'defined as the relative speed with which an innovation is adopted by members of a social system' (ibid). Apart from educating audio practitioners on the benefits of loudness normalisation within the system, the creation of an environment that will dissuade the use of hyper-compression largely depends upon the adoption of music streaming by the consumer. The claims made by the MLA rely heavily on loudness bias being removed from music

reproduction altogether which music streaming services that apply loudness normalisation can provide. Therefore, the adoption rate of these services is critical to these claims. According to the International Federation of the Phonographic Industry (IFPI - which 'represents the interests of the recording industry worldwide'), music streaming is 'emerging as the digital sector's main driver of growth' (IFPI 2016, p. 15). Global revenue from streaming services increased over a five-year period from 2011 to 2015 from US$600million to US$2.89billion, with a 45.2% increase in 2015 alone (*figure 8*) (ibid).

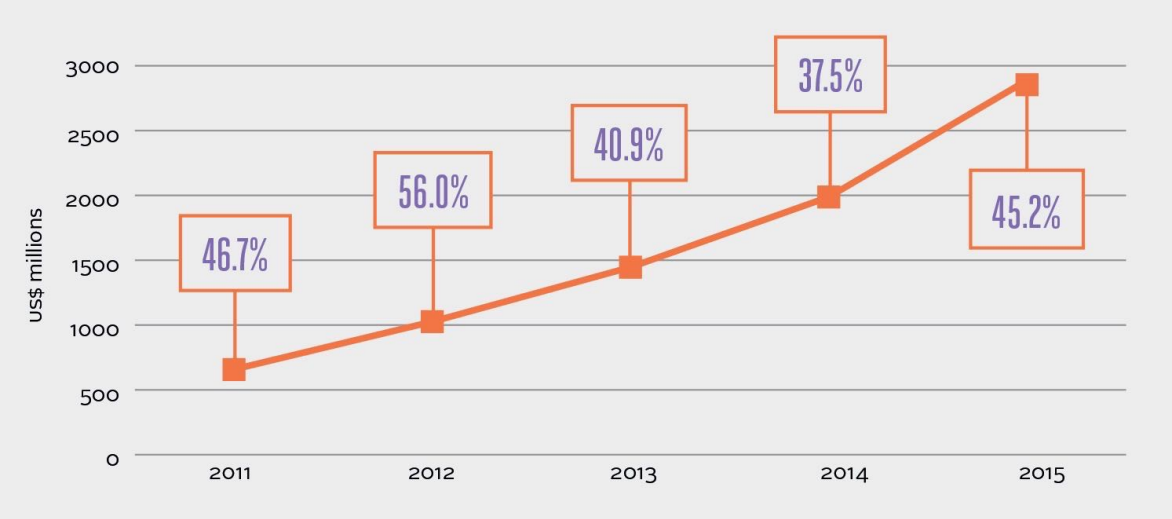

Figure 8: Streaming growth year on year, 2011-2015. Source: IFPI

Digital sales in general have shown a steady increase from US$4billion in 2008 to US$6.85billion in 2015 signalling the death of 'shopfront physical sales'; in 2015 'digital became the primary revenue stream, overtaking physical sales' for the first time (*figure 9*) (IFPI 2016, p. 8). Chief Executive of the IFPI, Frances Moore, argues that audiences have 'instant access, at any time and in any place, to a vast record collection of more than 43 million tracks (…) enabling the recording industry to reach markets that it could not monetise through physical retailing' (ibid, p. 5).

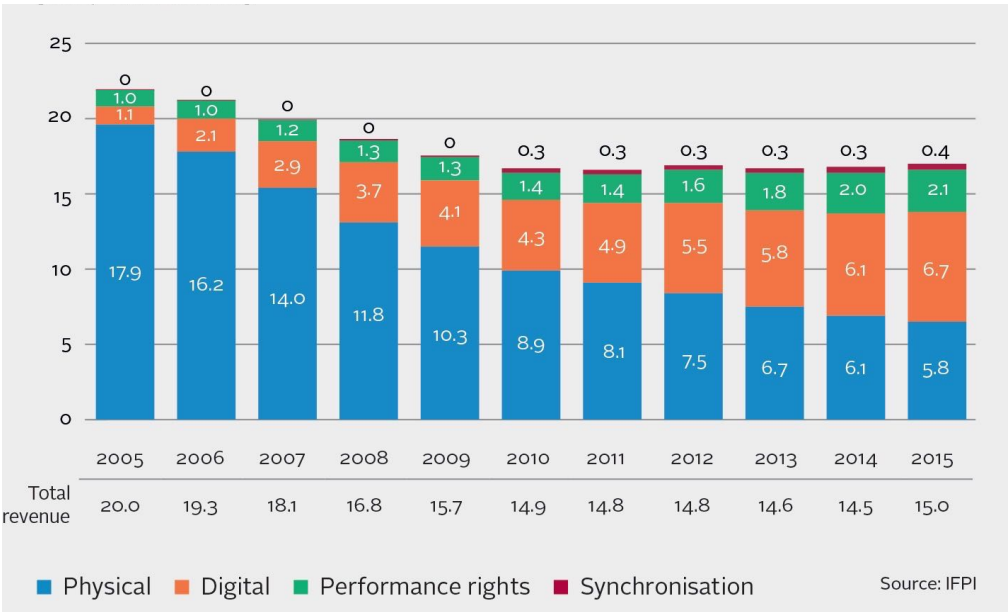

Figure 9: Global recorded music industry revenues 2005-20015, US$ billions. Source: IFPI

Within the category of digital sales which include both digital downloads and streaming services, streaming revenues 'have overtaken income from download sales in no fewer than 42 countries and account for 43 per cent of digital revenues globally' with 968 million users (*figure 10*) (IFPI 2016, p. 17). In contrast, revenue from digital downloads fell by '8% in 2014, an acceleration on the 2% drop in 2013', 'with double digit declines in most major markets' (IFPI 2015a, p. 21). According to another report by the IFPI in Sweden, the country where Spotify originated from, the uptake of streaming services has reached a remarkable '94% of the digital market' (IFPI 2015b). It is clear from these reports that music streaming is rapidly becoming the dominant platform for music delivery globally, and a trajectory that supports the claims made by the MLA. It could be argued that the speed in which music streaming is adopted could have a direct causal relationship to when the use of hyper-compression can be interrupted. Rogers asserts that 'innovations have an S-shape rate of adoption' with variation in the slope of the curve indicating the rate of adoption (2003, p. 23). From statistics detailed in this investigation we can approximate that the adoption rate is most likely below 50% of its potential as of 2015 (*figure 11*). It may require an adoption rate of

90% that firmly positions music streaming as the dominant music delivery platform. As Rogers argues, with the uncertainty surrounding innovation, time is the only way we can assess its success.

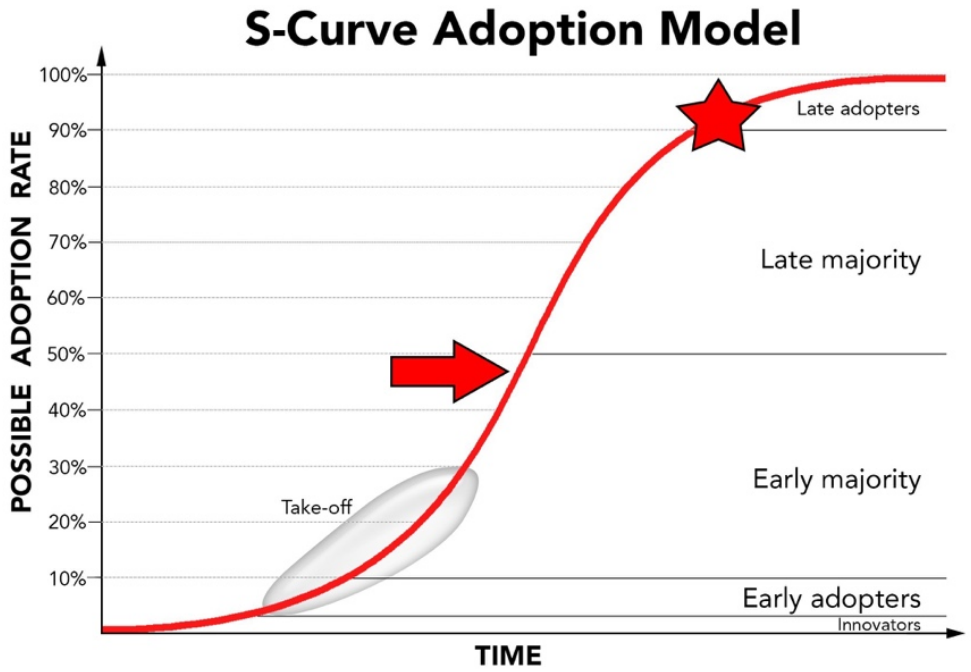

Figure 11: S-curve adoption model of music streaming adoption rate. The arrow indicates an approximation of adoption rate as of 2015. The star indicates a prediction where adoption rate will need to be for change within the social system.

Conclusion

Hyper-compression, a process whereby the dynamic range of sound recordings is dramatically reduced in order to increase perceived loudness on reproduction has been a contentious issue in audio production practices due to the undesirable artefacts the process introduces to the audio signal. However, the motivation for the use of hyper-compression in contemporary music has been driven by a psychoacoustic phenomenon by which listener's preferences are biased towards louder music when a comparison is made. Hyper-compression therefore has become a structural determinant within music production practices to allow

music recordings to operate within the market place alongside competitors. Loudness normalisation, a new technology originating from television broadcast, provides a method of regulating the loudness of program content, removing loudness discrepancies that may be potentially disruptive to the flow of content on television and radio. This innovation has also been found to be a useful application for file-based audio reproduction such as music streaming services and personal media players whereby music from different eras that contain large discrepancies in overall level can be managed to provide a more seamless listening experience. It has also been suggested that incorporating this technology into the audio reproduction setting may provide a solution to the proliferation of hyper-compression by removing loudness bias from the listening experience and therefore removing any motivation for the continued use of the process.

It is the contention of this paper that the introduction of loudness normalisation and its innovative application in both broadcast and music reproduction settings provides an effective solution to many of the longstanding problems concerning loudness. This innovation also poses a solution to the problem of the overuse of hyper-compression in music production, although as argued, loudness normalisation provides only a part of a potential solution. The solution requires the successful communication to the social system of the potential benefits of a loudness normalised reproduction paradigm, facilitated by what Rogers calls the 'opinion makers and change agents' (ibid, p. 24). In this case, the Music Loudness Alliance are performing their role as change agents via their activities in the AES community. Whether they are successful in garnering widespread support is unclear.

As part of the necessary contextual change mentioned throughout the paper which has the

potential to change the structure within the audio social system, this loudness normalised reproduction paradigm largely requires music streaming to become the dominant music delivery platform, disrupting the traditional model of physical sales; a model that propagated the use of hyper-compression. It is clear from the discussion within that what Rogers labels as the 'time dimension' element of diffusion, the trajectory of music streaming as a dominant delivery platform is highly probable yet uncertainty surrounds all innovation (2003, p. 20). At what point of adoption by the consumer, who form a significant part of the social system, will we see any evidence that the position hyper-compression plays within the social system has changed? There is no evidence of any change in the use of hyper-compression signalling that the adoption rate of music streaming has not yet reached that critical point.

References

AES (Audio Engineering Society) 2015, *Recommendations for Loudness of Audio Streaming and Network File Playback*, TD1004.1.12-10, technical document.

ATSC (Advanced Television Systems Committee) 2013, *Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television* (A/85:2013).

Camerer, F 2015, interview, Vienna, Austria, 21 October.

Camerer, F. 2012, *On the Way to Loudness Nirvana: Audio Levelling with EBU R-128*, EBU Technical review Q3, edited by Mike Meyer, European Broadcasting Union, Geneva, Switzerland.

EBU (European Broadcast Union) 2011a, *EBU Recommendation R-128*, Geneva, Switzerland. August.

EBU (European Broadcast Union) 2011b, *EBU Tech 3341 - Loudness Metering: 'EBU Mode' Metering to Supplement Loudness Normalisation in Accordance with EBU R-128*, Geneva, Switzerland, August.

EBU (European Broadcast Union) 2011c, *EBU Tech 3342 - Loudness Range: A Measure to Supplement Loudness Normalisation in Accordance with EBU R-128*, Geneva, Switzerland. August.

EBU (European Broadcast Union) 2011d, *EBU Tech 3343 V2.0 - Practical Guidelines for Production and Implementation in Accordance with EBU R-128*, Geneva, Switzerland. August.

EBU (European Broadcast Union) 2011e, *EBU Tech 3344 V1.1 - Practical Guidelines for Distribution Systems in Accordance with EBU R-128,* Geneva, Switzerland. October.

EBU (European Broadcast Union) 2012, *EBU Technology Fact-Sheet - Loudness V1.0*." Geneva, Switzerland.

Eggerton, J 2014, 'FCC Updates Calm Act to Further Quiet Commercial', *Broadcast and Cable*, Newbay Media, New York, USA, 6 April, viewed 12 February 2016. http://www.broadcastingcable.com/news/washington/fcc-updates-calm-act-further-quiet-commercials/131571

Fletcher, H & Munson, WA 1933, 'Loudness, its definition, measurement and calculation', *the Journal of the Acoustical Society of America*, vol. 5, pp. 82-108.

Free TV 2012, 'Measurement and Management of Loudness in Soundtracks for Television Broadcasting - operational practice OP- 59', Australia.

IFPI (International Federation of the Phonographic Industry) 2015a, *Digital Music Report*.

IFPI (International Federation of the Phonographic Industry) 2016, *Digital Music Report*.

IFPI (International Federation of the Phonographic Industry) 2015b, 'Sweden: A Market Transformed', viewed 21 January 2016. http://ifpi.org/sweden.php

ITU (International Telecommunication Union) 2012, *ITU-R BS.1770-3: Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level*.

Katz, Bob 2015, *interview*, New York, United States, 1 November.

Katz, Bob. "The Loudness War Has Been Won", Press release, 14/10/2013. http://www.digido.com/forum/announcement/id-6.html (Accessed: November 2014)

Ludwig, B 2015, *interview*, New York, United States, 30 October.

Milner, G 2009, *Perfecting sound forever—the story of recorded music*, Granta, London.

Powell, D 2009, 'Reducing the Volume of Ads', *Inside Science*, American Institute of Physics, Maryland, United States, 17 September, viewed 23 February 2016. https://www.insidescience.org/content/reducing-volume-tv-ads/1324

Rogers, E M 2003, *Diffusion of Innovations*, fifth edition, Free Press, Simon and Shuster Inc., New York.

Ronan, M, Sazdov, R & Ward, N 2014b, 'Factors influencing listener preference for dynamic range compression', *proceedings of the 137th Audio Engineering Society Convention,* October, New York.

Taylor, RW and Martens, WL 2014, 'Hyper-compression in music production: listener preferences on dynamic range', *proceedings of the 136th Convention of the Audio Engineering Society*, April, Berlin.

Thiele, N 2005, 'Some thoughts on the dynamics of reproduced sound', *The Journal of the Audio Engineering Society*, vol. 53, no. 1/2, pp. 130-132, January/February.

Vickers, E 2010, 'The Loudness War: Background, Speculation and Recommendations', *proceedings of the 129th Convention of the Audio Engineering Society*, November, San Francisco.

Weymouth, D 2012, *The loudness war: a game and market theory analysis*, essay, GRIN, Munich.