Adaptive Control of Amplitude Distortion Effects

Brecht De Man[1], Joshua D. Reiss[1]

[1]*Centre for Digital Music, Queen Mary University of London, London E1 4NS, U.K.*

Correspondence should be addressed to Brecht De Man (`b.deman@qmul.ac.uk`)

ABSTRACT

An adaptive amplitude distortion effect for audio production is introduced. We consider a versatile model of multiband waveshaping with a limited set of intuitive parameters. We propose a method for adaptive scaling of the characteristic amplitude distortion transfer function. A number of methods to automate other distortion parameters are explored, such as automatic balance between clean and distorted signals for single band and multiband distortion, make-up gain and adaptive anti-aliasing. Automation formulas are generalised for several other common transfer functions. A formal perceptual evaluation of the automation algorithms is conducted, validating the approach and identifying shortcomings and particularities. Finally, we discuss implementational aspects and possible future directions.

1. INTRODUCTION

Being an important characteristic of many analog audio devices such as microphones, amplifiers, mixing consoles, effect processors, and analog tape, distortion has always been an important effect in music production. Though a nuisance for most electronics designers, a warm, subtle distortion is often one of the most sought-after features in audio gear, and no less since the advent of digital audio production [1].

In this text, distortion refers to applying any type of non-linear transfer function on a waveform (referred to as amplitude distortion or waveshaping, as opposed to phase or time-domain distortion [2–5]). In the digital domain, this can be any non-linear function defined, for simplicity, from $[-1, 1]$ (input) to $[-1, 1]$ (output). As such, it is a very simple, 'low-cost' audio effect as it can be implemented through calculation of a simple mathematical function or a lookup table [2].

In general, sound engineers, guitarists and other musicians tend to feel digital distortion sounds inferior to the 'original' analog gear [6]. Much work has been done on the emulation of analog, non-linear devices [4, 6–10]. Distortion as a creative, sound shaping effect owes its popularity to rock guitarists, but used gently, distortion effects can alter timbre and dynamics in sophisticated and subtle ways for all kinds of sources [11]. Beyond its creative, sound-shaping use, harmonic distortion can be of practical importance by increasing the audibility of sources and/or as an alternative for equalisation [3]. Carefully applied distortion can even enhance the perceived quality and presence of audio signals [11]. A soft clipping curve or similar can also be applied to decrease the dynamic range, and make the perceived loudness greater for the same peak level [11].

As the amount of distortion usually increases with increasing input signal level [2, 12], a static distortion device needs careful setting of the gain and/or threshold controls to apply the right amount of distortion. This also means that static settings inevitably introduce varying amounts of distortion for different input signal levels. This may be desirable in some cases (after all, this is the expected behaviour of the intrinsic saturation induced by analog devices such as amplifiers) but can be detrimental in situations where a constant degree of distortion is desired, leaving soft parts virtually unprocessed and/or louder parts heavily damaged. It also requires time and effort of an audio engineer to tailor the characteristic threshold of the distortion transfer function to the level of the input source.

In this paper, we describe a distortion effect that scales the transfer function in real-time proportional with a moving average of the signal's level. As such, it is a member of the Adaptive Digital Audio Effects (A-DAFx) in the sense meant by [13].

Adaptive effects are especially important in the context of automatic mixing systems [14]. In these systems, little or no user interaction is desired, meaning that all parameters need to be set automatically by the system itself based on features extracted from the incoming audio or other available knowledge. In this work, we add amplitude distortion to the growing set of automatic audio effects for music production such as automatic panning [15–17], balancing [18, 19], equalising [20, 21], and dynamic range compression [22]. However, the considered effect derives part of its parameters from the audio input in an autonomous way, whereas other parameters (the amount and type of distortion) are still for the user to define, to be set as a constant characteristic to that system, or to be based on other (semantic) data, extracted by the mixing system or provided by the end user. As such, it can be readily incorporated in autonomous systems where processing parameters are derived from song or instrument data in conjunction with sets of semantic, user-definable rules [23].

As the effect needs to act autonomously (i.e. without human interaction) in many situations, it is undesirable to control the amount of distortion by an input gain control, which increases the effective level and changes the relative level of the different bands in the multiband case. Rather, we control the parameters of the device's transfer function, manipulating for instance its clipping point, while compensating automatically for the loss in loudness or energy induced by clipping the signal's peaks. Another reason for scaling the characteristic transfer function rather than the input gain to control the amount of distortion is that the latter introduces a level difference in 'before/after' tests, making careful comparison impossible.

We consider a multiband distortion device with different amounts and transfer curves per band to be able to match typical analog devices that distort different frequencies differently [24]. Many exciter implementations only distort high frequencies, which also calls for a multiband implementation [1]. Another motivation for the multiband approach is the ability to increase audibility of low frequencies that can not be heard adequately on some playback systems, by creating harmonics of these frequencies in particular. Finally, for polyphonic instruments, ensembles or any broadband sound for that matter, different notes, tones or other acoustic events can occur at once or overlap, which sometimes leads to brutal intermodulation distortion [25]. Distorting different bands, where they do not interact, can reduce these effects and for example allow for a large amount of distortion with only little intermodulation between different bands.

Throughout this paper, we investigate the possibility of automating multiband amplitude distortion such that

- no human interaction is required to set an appropriate threshold or other characteristic transfer function parameter, maintaining an amount of distortion (non-linearity) independent of input level;
- no excess or deficit of this distortion emerges at the onset of a louder or softer audio fragment, with an imperceptible transition;
- aliasing artefacts are diminished down to an acceptable standard while not using a higher upsampling ratio than necessary, based on knowledge of the highest significant frequency of the input;
- the level does not change with regard to the original signal level.

Specifically, we demonstrate the first two aspects through perceptual evaluation.

In the following sections, we describe one particularly versatile function based on but not limited to the traditional hard and soft clipping functions, and propose several parameter automation methods. We then translate these techniques to other meaningful amplitude distortion curves. The approach towards automatic transfer function scaling is validated through a formal listening test, and results are discussed. Finally, we summarise our conclusions and outline future perspectives.

2. DISTORTION MODEL

A variety of transfer functions can be used for amplitude distortion (see Section 3.5 for translation of the principles presented here to other transfer functions), but in this section we will describe a very flexible, 'standard' curve. This will then allow us to propose and investigate adaptation methods using this simple but powerful transfer function parametrisation as an example. A block diagram outlining the complete system is shown in Figure 1. The automated effect is implemented as a VST plugin, with the interface as in Figure 2.

In the simplest case, the signal above the threshold value T (below $-T$) is clipped to T ($-T$). Similar to a dynamic range compressor, a soft 'knee' transition between

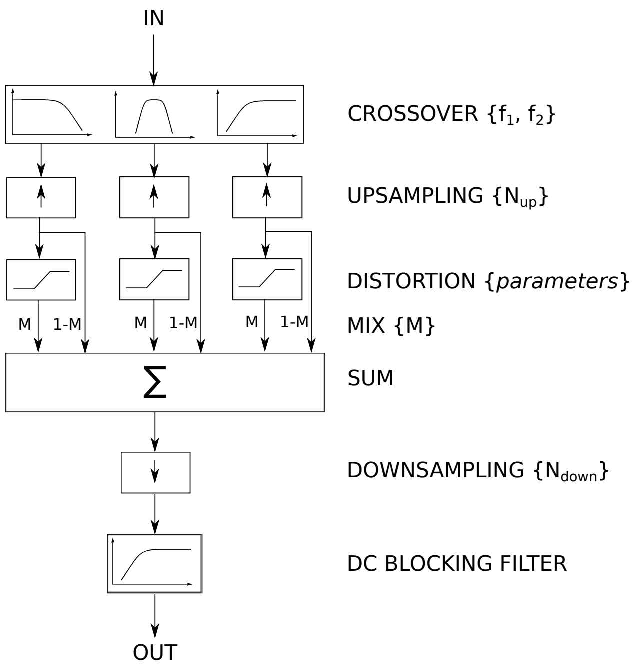


Fig. 1: Diagram of the complete system.

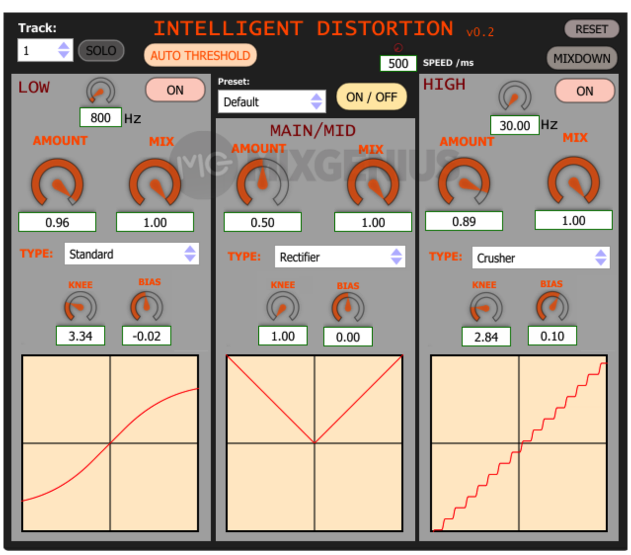


Fig. 2: VST plugin interface (developed in collaboration with Mixgenius Inc.).

the linear and compressed/clipped part is presented as an alternative to 'hard clipping', where the transition is abrupt [26]. As such, the signal only clips for (absolute) values above this transition region (the 'knee', i.e. $|x| > T \cdot \sqrt{K}$). where x is the (clean) input signal, T the transfer function's threshold, and K the knee width. Note that we use linear values (x between -1 and 1, T between 0 and 1, $K > 1$ for soft knee and $K = 1$ for hard knee, i.e. no transition region at all) following common practices [11, 27], rather than values in decibel as is customary for dynamic range compressor formulas [26]. For this reason, the knee, which in decibel is defined as the region from $T_{dB} - K_{dB}/2$ until $T_{dB} + K_{dB}/2$, is defined as the region from $T/\sqrt{K}$ to $T \cdot \sqrt{K}$ in the linear domain.

For the knee transition (between $|x| = T/\sqrt{K}$ and $|x| = T \cdot \sqrt{K}$), we use a cubic Hermite spline to connect the linear ($f(x) = x, |x| \leq T/\sqrt{K}$) and the clipped part ($f(x) = \mathrm{sgn}(x) \cdot T, |x| \geq T \cdot \sqrt{K}$), which yields (for $T/\sqrt{K} < |x| < T \cdot \sqrt{K}$):

$$f(x) = \mathrm{sgn}(x) \frac{K\,|x|^3 - T(2K^{3/2} + K + 2K^{1/2})\,|x|^2 + T^2(K^2 + 2K^{3/2} + 4K)\,|x| - T^3(2K^{1/2} + 1)}{(K^2 + 2K^{3/2} - 2K^{1/2} - 1) \cdot T^2} \quad (1)$$

Bringing a bias 'voltage' B into the equation, we can create curves where the positive end clips 'sooner' than the negative end, or vice versa. Finally, a mix parameter M enables a straightforward balance of the distorted and original signal, enabling parallel processing.

This set of parameters provides an intuitive way of determining a variety of transfer functions, including elementary ones like hard and soft clipping, as well as tube-like overdrive. To facilitate the setting of parameters, we provide a visualisation of the transfer function on the interface of the VST plugin (see Figure 2).

3. PARAMETER AUTOMATION

3.1. Adaptive transfer function scaling

In order to obtain an equal amount of distortion (for a given set of parameters) for signals of varying level, we implement an 'auto-threshold' function, similar to the automatic threshold in automatic compressor designs [22].

To ensure that the same amount of distortion is applied for different levels, we scale the threshold with the square root of the moving average of the square of the signal value (moving RMS). This means that, by definition, changing the signal level by a certain amount changes the threshold by that same amount.

Instead of threshold T, which we will automate, we give the user control over the the relative threshold T_{rel}, which relates to the threshold as follows:

$$T = T_{rel} \cdot L_{RMS} \tag{2}$$

where L_{RMS} is the moving RMS level and T_{rel} is the threshold relative to the smoothed RMS level. Note that this relative threshold, unlike the effective threshold T, is not necessarily less than 1.

Smoothing the level measure is performed through an exponential moving average (EMA):

$$L_{RMS}[n] = \sqrt{(1-\alpha) \cdot x^2[n] + \alpha \cdot L_{RMS}^2[n-1]} \tag{3}$$

where $x[n]$ is the value of sample n, and

$$\alpha = \exp\left(\frac{-1}{\tau \cdot f_s}\right) \tag{4}$$

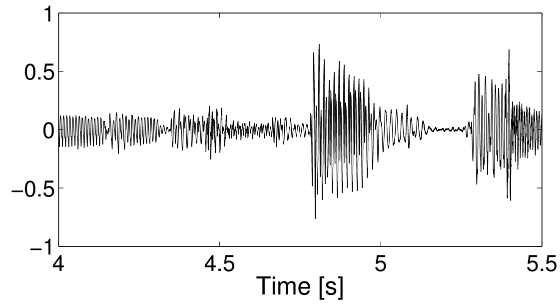
where τ is the characteristic time constant in seconds, and f_s is the sampling rate [26].

A possibly advantageous effect of this is that the dynamic range is left unharmed, even for very high degrees of distortion. As an example, static hard clipping with a very low threshold could bring down the level of a rather loud fragment to the level of a subsequent soft part of the same source, as shown in Figure 3(b). When made adaptive as above, however, the long-term dynamics will be maintained as not only the same distortion will be applied, but the relative level differences will also be maintained, as opposed to traditional distortion devices. Note that, depending on the value of the time constant, relatively short-term dynamics will still be reduced.

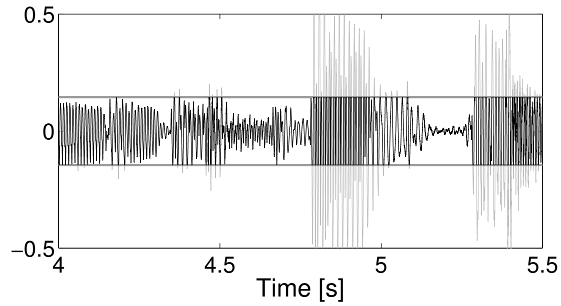
Similarly, the bias should be proportional to the same factor, which effectively means that the whole transfer function is scaled with the smoothed RMS level. Since the knee K and mix M are 'relative' measures, they do not have to be scaled.

3.2. Adaptive transfer function scaling - improved

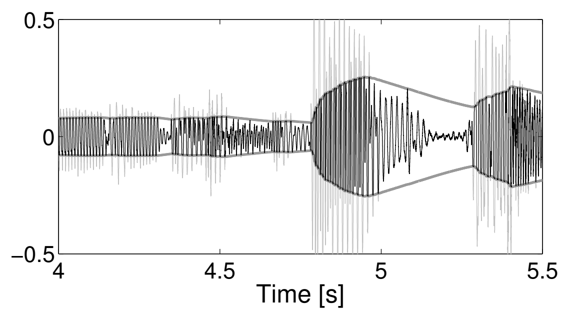
As evident from Figure 3(c), the lag inevitably introduced by the exponential moving average causes the



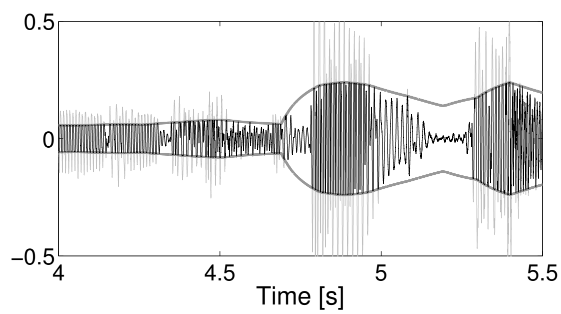
(a) Clean signal



(b) Distorted signal: static threshold



(c) Distorted signal: adaptive threshold



(d) Distorted signal: adaptive threshold with lookahead function and peak detection mechanism

Fig. 3: Fragment of bass guitar recording with sudden level change. In (b) through (d), the light grey signal is the original, clean signal, while the black signal is the resulting signal after symmetric hard clipping distortion with the threshold shown in dark grey.

threshold to react late to onsets of parts with a significantly different level. Consequently, a quiet signal followed by a louder one would temporarily be distorted more than desired, and vice versa. This can be compensated by choosing ever lower time constants, causing the threshold, and hence the amount of distortion introduced, to change much faster. Another solution to this problem is the introduction of a lookahead function (delayed output while the input is considered a number of samples in advance) in combination with a peak detection into the previously described, sample-based system. Equation 3 then becomes:

$$L_{rms}[n] = \sqrt{\begin{array}{l}(1-\alpha)\cdot\max\left(x^2[n], x^2[n+1], ..., x^2[n+L]\right)\\ \quad +\ \alpha\cdot L_{rms}^2[n-1]\end{array}} \tag{5}$$

The effect of using such a lookahead buffer (with characteristic adaptive time constant $\tau = 200$ ms, the corresponding EMA coefficient α, and lookahead buffer size $L \cdot f_s = 100$ ms) is shown in 3(d).

3.3. Anti-aliasing

Like other non-linear effects, amplitude distortion introduces harmonics at frequencies that potentially were not present in the original signal. This means it may be necessary to perform sufficient upsampling before applying distortion to avoid audible aliasing artefacts [29]. The extra frequency components introduced by aliasing are almost always inharmonic, and therefore deteriorate the perceived sound quality [11]. For this reason, we apply an interpolation filter to the signal before applying any kind of processing that introduces new harmonics. Afterwards, if an audio output at the original sampling frequency is desired, a decimation filter should be applied. Naturally, a higher upsampling ratio is more computationally expensive.

The effect of upsampling before applying harmonic distortion is shown in Figure 4. The up- and downsampling in these figures as well as in the MATLAB implementation of the presented distortion effect is performed by the MATLAB function `resample`.

Depending on the implementation, it may be more or less computationally expensive to upsample the filtered bands (as shown in Figure 1), or to upsample the input signal first and only then go through the crossover filter stage. In some cases it may be desirable to only upsample the highest band if it is expected that lower bands will not produce audible aliasing artefacts. The lower bands will then contain less samples to be processed allowing the distortion to be applied more quickly.

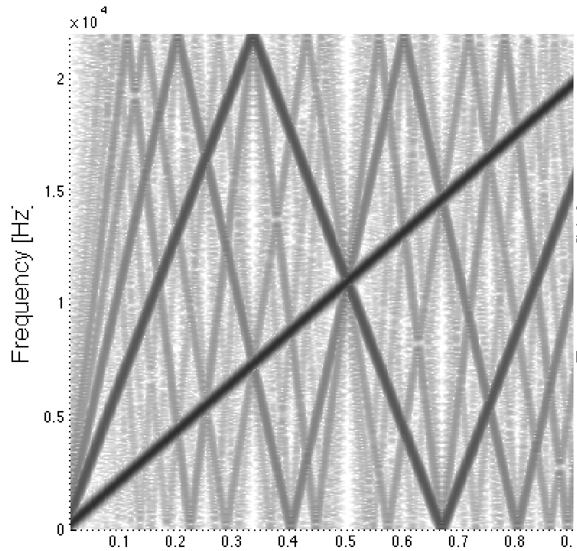
Knowing the signal's or band's approximate upper frequency limit and estimating the number of harmonics that should be below the Nyquist frequency, an adaptive upsampling ratio can be found to save computational power while ensuring a sufficiently low perceptual difference. For a signal where the highest significant signal component is f, the desired number of new harmonics to fall below the new Nyquist frequency is N, and the original sampling rate is f_s, the necessary upsampling ratio can be found as follows:

$$R = \left\lceil \frac{2Nf}{f_s} \right\rceil \tag{6}$$

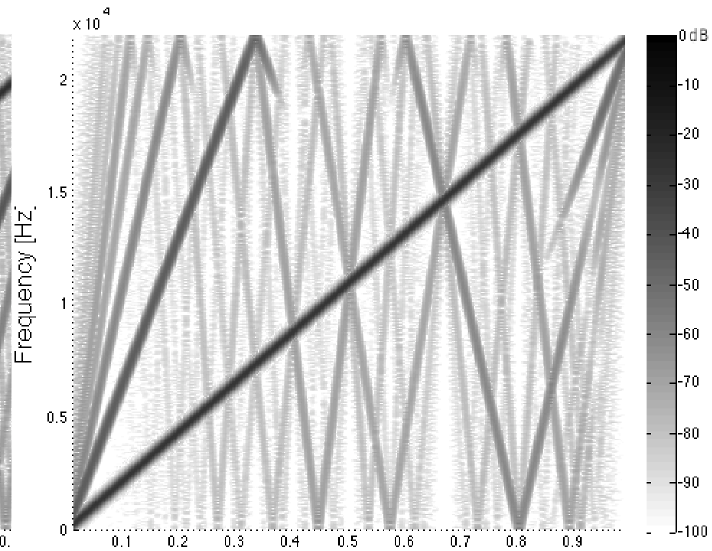
Further automation of the upsampling ratios could be achieved by making them dependent on the type of distortion (more specifically the relative levels of the respective harmonics), user input (the desired maximum amount of computation and/or minimum degree of quality) and possibly the time-varying spectrum of the input signal, if this information is extracted or provided. In the special case of polynomial transfer functions of n^{th} order, new frequency content is only generated up to the n^{th} harmonic [5], making even simpler, real-time adaptive upsampling possible [9].

3.4. Further automation

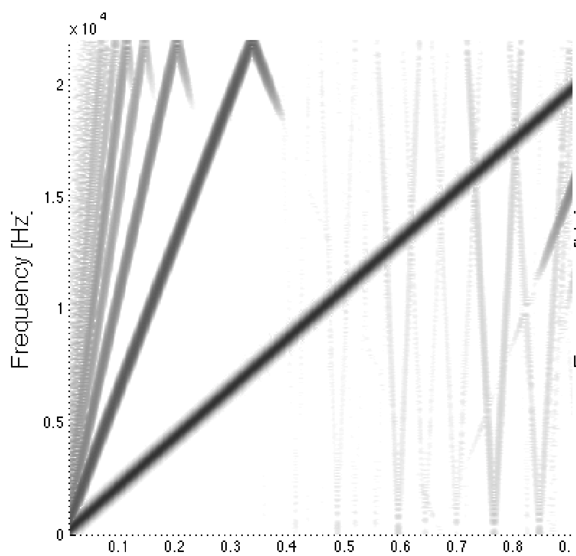
To address the possible change in loudness induced by the distortion, another automation mechanism is provided. The level or loudness is measured of each filtered signal component before distortion (this can be the RMS level value or a more sophisticated loudness measure), and then again after applying distortion. The distorted signal is then multiplied with a smoothed version of the ratio of these values L_{in}/L_{out} before being mixed with the clean input signal. The exponential moving average of this ratio is based on another characteristic time constant, τ_L, as in Equation 3. Depending on the accuracy of the value, this ensures that there is no perceivable difference in loudness or level when the device is switched on or off, or when the mix parameter M takes another value. More importantly perhaps, it avoids a level drop of the distorted signal when the threshold is brought down, due to 'chopping' off the signal's peaks to a degree where



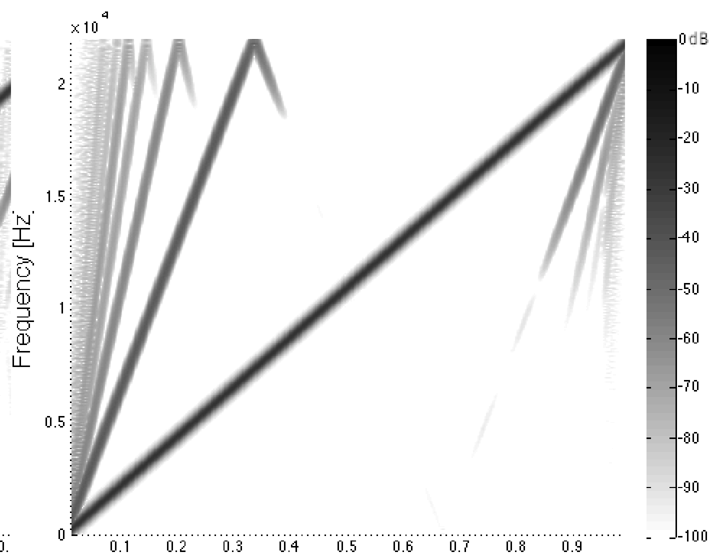
(a) No upsampling



(b) Upsampling ratio 2



(c) Upsampling ratio 8



(d) Upsampling ratio 32

Fig. 4: The above spectrograms show a linear sine sweep from 0 Hz to $f_s/2 = 22.05$ kHz with amplitude 1 being upsampled and low-pass filtered; hard clipped at $T = 0.5$; and low-pass filtered and downsampled again. As expected, the aliasing is especially problematic at high frequencies, and decreases (but never completely disappears) with an increasing upsampling factor.

little remains of the original signal. In the case of multi-band distortion, it prevents the bands from changing their relative loudness (e.g. a distorted, louder high band), as the absolute loudness or level of every single band is maintained.

A last automation mechanism is to impose a certain ratio of the energy of the high band versus the energy of the rest of the signal, and mix in more or less high band distortion to achieve this target ratio and avoid dullness or piercing highs. In this case, the energy of a high-pass filtered signal above $2 \cdot f_2$ is compared to the total signal energy (where f_2 is the cutoff frequency of the high-pass filter, see Figure 1), and the distortion of the high-pass filtered signal x_{HPF} is balanced accordingly. Note that we consider the energy of the frequencies above $2 \cdot f_2$, as harmonic distortion of the high band starting at f_2 will have its the second harmonic at $2 \cdot f_2$ or higher. As such, we have designed an automatic exciter, ensuring a consistent amount of brightness in the processed signal.

Note that although many intrinsic distortion parameters can be automated through the above measures, there are still at least as much parameters that need tweaking (time constants τ and τ_L, relative threshold T_{rel}, the target high-to-low energy ratio, ...). However, these measures can be set to appropriate values once, possibly saved as a preset, and are by definition independent of the signal level from that point onwards, making it safe to leave them unchanged in an autonomous mixing situation.

3.5. Other transfer functions

Virtually any amplitude distortion transfer function is suitable for threshold automation: the underlying principle is to scale the input variable x with $1/L_{RMS}$ to ensure that similar signals with different levels (captured in L_{RMS} for sufficiently long-term dynamics) have the same relative amount of distortion. Similarly, the input should be scaled with the relative threshold T_{rel} to set the relation between the characteristic threshold of the function (depending on its definition, this is the amplitude where a certain deviation from a linear curve is reached) and the moving average of the level L_{RMS}.

4. PERCEPTUAL EVALUATION

4.1. Test design

To evaluate the perceptual effect and validate the approach taken towards adaptive amplitude distortion independent of input level, we conducted a listening study with $N = 12$ participants where the approaches described in Section 3.1 and 3.2 are assessed alongside a clean (reference) and statically distorted sample (anchor). We chose a MUSHRA-style test design because of the number and type of stimuli to be assessed [28].

As pilot tests showed that subtle amounts of distortion, though perceptually pleasing and reminiscent of analogue devices, were very hard to assess, we chose to use clearly distorted samples for this listening test. For the sake of generality, a number of very different sources were used.

To minimise audible aliasing, the samples were stored with a 96 kHz sampling rate. We use real-world, commercial grade stems of under 10 seconds length, with an artificial, instantaneous 10 dB boost in the second half of the fragment. Four different source samples were used: a bass guitar, drums, electric piano and vocal recording. The samples were labeled in a random order, to avoid bias from the suggested order or from remembering ratings of the previous instrument. In contrast to a traditional MUSHRA test, the participants were not required to put any of the samples at maximum value, so as not to restrict their answers. To approach a fair comparison, all threshold curves for the same instrument (static, adaptive and adaptive with lookahead) have the same average.

The subjects were asked to rate the different samples on the following aspects:

- Consistency of distortion throughout sample
- Amount of distortion in soft part
- Amount of distortion in loud part

They were also able and encouraged to make comments on any of the samples, as well as general comments.

4.2. Results and discussion

Of the 12 participants, 3 were taken out in the displayed results due to very different ratings for the clean reference case (low consistency and high distortion), presumably because of poor understanding of the test assignment.

Figure 5 shows the subjects' ratings for consistency, as well as the difference of distortion between the first and second half of the samples. From these results, it is clear that there is a high perceptual consistency throughout each sample for the adaptive distortion case. Similarly,

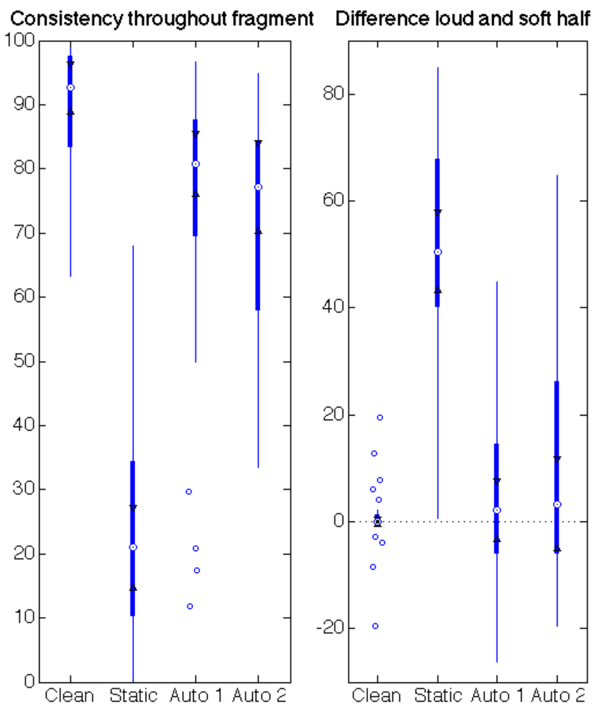


Fig. 5: Results from the listening test (N=9). Left are the ratings on 'Consistency of distortion', right are the differences between the perceived amount of distortion of the loud and the soft half of each fragment. Following standard practice, the median is represented by the dot, the whiskers extend from minimum until maximum (not considered outliers), the edges of the box are the 25th and 75th percentiles, and the circles represent the outliers. If the regions marked by the black triangles do not overlap, their medians differ significantly at the 5% significance level.

the difference in perceived distortion level between the low and high volume part of the fragment is close to zero on average, although with significantly greater variance than for the clean reference case. There is no significant difference between the ratings of the two adaptive cases. Although the perceived level of distortion was rated higher on average in the first case (Auto 1), this is certainly not a general trend when looking at the ratings for each instrument individually.

In 33 out of 36 cases, the static distortion was found to be less consistent in terms of perceived distortion than either of the adaptive ones. In again 33 out of 36 cases (not necessarily the same instruments and subjects), the perceived difference in terms of distortion between the loud and soft part was rated higher in the static case than it was in the adaptive cases.

When looking at the various instruments individually, the same trends are evident. Participants who commented on the ease of hearing differences in distortion noted that the electric piano recordings were particularly easy, whereas the vocal was found to be rather challenging. Opinions about the drum recording were mixed.

On five separate occasions, subjects commented they weren't able to perceive much difference between the two adaptive approaches, if any. For this reason, the variance may be disproportionally large for these cases, as the ratings were often rather far apart, even when little difference was reportedly perceived. It is possible that if the difference between the adaptive and improved adaptive distortion were bigger, a clearer trend would have been visible. Note that few participants commented extensively.

Although the degree of nonlinearity (level of harmonics) is mathematically the same regardless of the level after a transition period, it is interesting to note that there is a slight tendency to rate the louder part as more distorted in the adaptive cases (see also Figure 5).

On the whole, it can be inferred that the adaptive distortion does an adequate job equalising the (perceived) amount of distortion for different signal levels. There seems to be little real difference between the two algorithms under test.

5. CONCLUSION

We described a method of automating amplitude distortion parameters to achieve a consistent amount of distortion for signals of different levels, obviating the need for a human operator to set the effect's threshold or scale across different channels or for different sections of a song. We also explored several other automation strategies for amplitude distortion effects, to maintain a target spectrum, loudness, level, and a maximum level of aliasing artefacts. Furthermore, a general method was proposed to create adaptive distortion based on any kind of transfer function, and demonstrated it with other transfer function types.

We evaluated instances of two adaptive transfer function scaling methods, and found that they maintained distortion amount adequately. As there was no evidence of one method outperforming the other, it can be assumed the simple and less computationally expensive algorithm would suffice for most applications.

6. FUTURE PERSPECTIVE

The evaluation of the other proposed automation methods warrants investigation.

Further research will be needed to map the relative threshold control in such a way that it corresponds with a perceived amount of distortion regardless of the transfer function type. Perceptual distortion measures such as a filtered version of the signal deviation [29] can be instrumental in this matter.

More data and a demonstration of the effect can be found on `www.brechtdeman.com/research.html`.

7. ACKNOWLEDGEMENTS

The authors would like to thank Stuart Mansbridge for his help with the VST implementation.

8. REFERENCES

[1] U. Zölzer, Editor, *DAFX: Digital Audio Effects*, 2nd ed., J. Wiley & Sons, 2011.

[2] J. Timoney, V. Lazzarini, B. Carty and J. Pekonen, "Phase and amplitude distortion methods for digital synthesis of classic analog waveforms," presented at the 126th Convention of the Audio Engineering Society, May 2009.

[3] A. U. Case, *Sound FX*, Taylor & Francis, 2012.

[4] J. Timoney, V. Lazzarini, A. Gibney and J. Pekonen, "Digital emulation of distortion effects by wave and phase shaping methods," *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, September 2010.

[5] M. Le Brun, "Digital waveshaping synthesis," *Journal of the Audio Engineering Society*, vol. 27, pp. 250-266, April 1979.

[6] D. T. Yeh, J. S. Abel and Julius O. Smith, "Simplified, physically-informed models of distortion and overdrive guitar effects pedals," *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*, September 2007.

[7] D. T. Yeh and J. O. Smith, "Discretization of the 59 Fender Bassman tone stack," *Proceedings of the International Conference on Digital Audio Effects (DAFx-06)*, pp. 1-6, September 2006.

[8] A. Huovilainen, "Nonlinear Digital Implementation of the Moog Ladder Filter," *Proceedings of the International Conference on Digital Audio Effects (DAFx-04)*, October 2004.

[9] J. Schimmel, "Using nonlinear amplifier simulation in dynamic range controllers," *Proceedings of the 6th International Conference on Digital Audio Effects (DAFX-03)*, September 2003.

[10] David T. Yeh, Balzs Bank and Matti Karjalainen, "Nonlinear modeling of a guitar loudspeaker cabinet," *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, September 2008.

[11] S. Enderby and Z. Baracskai, "Harmonic instability of digital soft clipping algorithms," *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, September 2012.

[12] J. Vanderkooy and K. B. Krauel, "Another view of distortion perception," presented at the 133rd Convention of the Audio Engineering Society, October 2012.

[13] V. Verfaille, U. Zölzer and D. Arfib, "Adaptive Digital Audio Effects (A-DAFx): A New Class of Sound Transformations," *IEEE Transactions on Audio, Speech, and Language Processing*, 2006.

[14] J. D. Reiss, "Intelligent systems for mixing multichannel audio," presented at the 17th International Conference on Digital Signal Processing (DSP), July 2011.

[15] S. Mansbridge, S. Finn and J. D. Reiss, "An Autonomous System for Multi-track Stereo Pan Positioning," presented at the 133rd Convention of the Audio Engineering Society, October 2012.

[16] E. Perez-Gonzalez and J. D. Reiss, "A real-time semiautonomous audio panning system for music mixing," *EURASIP Journal on Advances in Signal Processing*, May 2010.

[17] E. Perez-Gonzalez and J. D. Reiss, "Automatic Mixing: Live Downmixing Stereo Panner," *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-10)*, September 2007.

[18] S. Mansbridge, S. Finn and J. D. Reiss, "Implementation and evaluation of autonomous multi-track fader control," presented at the 132nd Convention of the Audio Engineering Society, April 2012.

[19] E. Perez-Gonzalez and J. D. Reiss, "Automatic gain and fader control for live mixing," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, October 2009.

[20] E. Perez-Gonzalez and J. D. Reiss, "Automatic equalization of multi-channel audio using cross-adaptive methods," presented at the 127th Convention of the Audio Engineering Society, October 2009.

[21] E. Perez-Gonzalez and J. D. Reiss, "Improved Control for selective minimization of masking using inter-channel dependency effects," *Proceedings of the 11th International Conference on Digital Audio Effects*, September 2008.

[22] J. A. Maddams, S. Finn and J. D. Reiss, "An autonomous method for multi-track dynamic range compression," *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, September 2012.

[23] B. De Man and J. D. Reiss, "A Knowledge-Engineered Autonomous Mixing System," presented at the 135th Convention of the Audio Engineering Society

[24] R.-E. Gaskell, P. E. Gaskell, and G. Massenburg, "Distortions in audio op-amps and their effect on listener perception of character and quality," presented at the 131st Convention of the Audio Engineering Society, October 2011.

[25] M. Zaunschirm, J. D. Reiss and A. Klapuri, "A sub-band approach to musical transient modification," *Computer Music Journal*, Vol. 36 (2), Summer 2012.

[26] D. Giannoulis, M. Massberg and J. D. Reiss, "Digital Dynamic Range Compressor Design—A Tutorial and Analysis," *Journal of the Audio Engineering Society*, Vol. 60 (6), pp. 399-408, 2012.

[27] D. T. Yeh, *Digital Implementation of Musical Distortion Circuits by Analysis and Simulation*, Ph.D. dissertation, Stanford University, June 2009.

[28] B. De Man and J. D. Reiss, "A Pairwise and Multiple Stimuli Approach to Perceptual Evaluation of Microphone Types," presented at the 134th Convention of the Audio Engineering Society, May 2013.

[29] L. Vetter, M. J. Terrell, A. J. R. Simpson and A. McPherson, "Estimation of Overdrive in Music Signals," presented at the 134th Convention of the Audio Engineering Society, May 2013.