Measuring a Portable Audio Device's Response to Excessive Sound Levels

Paul Kendrick, Stephen Groves-Kirkby, Iain Jackson, Trevor Cox, Bruno Fazenda

July 5th 2013
The Good Recording Project
Acoustics Research Centre
University of Salford
http://www.goodrecording.net
p.kendrick@salford.ac.uk

1. Summary

This is a report detailing how a series of measurements were carried out to quantify the non-linear performance of a series of portable audio recorders. The performance of a series of common devices is quantified including the; Cannon 550D, Edirol r44, Neumann U87ai via Focusrite 2i4, shure SM57 via Focusrite 2i4, Zoom H2, Zoom H4, Google Nexus 4, iphone and a sony camcorder (vx2000). Parameters relating to attack release and thresholds of any dynamics processing present are reported, and the devices are probed fort any non-linear behaviour such as clipping, if any non-linear behaviour is identified it is parameterised by extracting the gain curve which relates input level to output level.

2. Introduction

Consumer devices capable of recording audio such as mobile phones, video cameras, dictaphones, laptops or tablets, often employ some form of gain control to prevent excessive levels resulting in distortion where the peaks of the waveform are clipped. Gain control systems may respond instantaneously, employing a nonlinear gain curve softening the clipped peaks; reducing the level of harmonic distortion, or employ attack and release transition periods, which reduces the quality degradation by applying dynamic gain adjustments. This report details a method to "blindly" analyse the performance of devices and define the attack, release, compression ratio and threshold parameters for dynamic range compression and gain curves for nonlinear distortions.

A number of previous authors have attempted to answer similar problems (Bitzer, Schmidt, & Simmer, 2006; Cabot, 1987; Schiffner & Betz, 1996). Generally these involve passing test signals through the device designed to cause the compressor to alter the gain of the system. This is not possible with portable audio recorders as simultaneous input-output measurement is not possible for end-users. As the portable recorders are all clocked using internal clock signal, synchronisation can be problematic. A similar method to (Bitzer et al., 2006) is employed in this work but with a number of modifications which address these problems.

This report first introduces the methodology in section 3 and section 4 details some results from a number of common portable recording devices.

3. Methodology

To capture the nonlinear dynamic response of each of these devices a signal (Figure 1) was generated consisting of a 0.5 s pink noise burst appended by a quieter (-20dB) 1 s pink noise burst. The louder first pulse is intended to overload the system or activate the gain control system. The quieter second pulse is to allow the level to drop below the threshold and capture the release behaviour of the system.

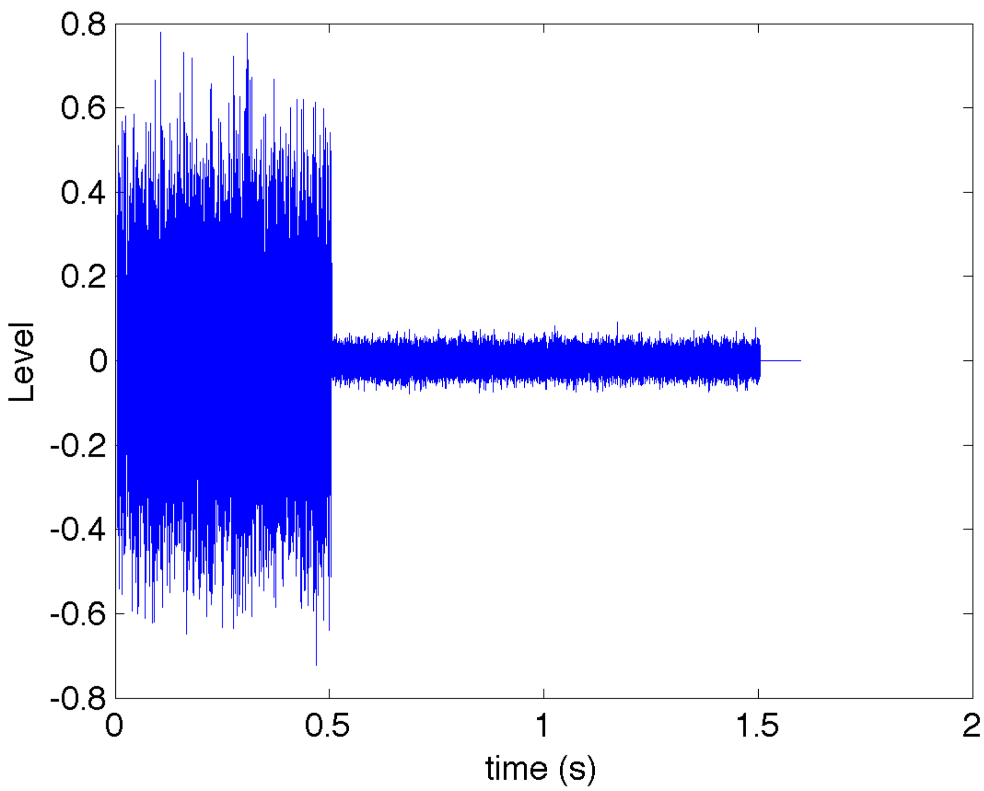


Figure 1. Pulse used to capture dynamic gain control systems

The test signal was presented to each device using a Genelec 1029A loudspeaker in an anechoic chamber at a distance of 0.5 m. The system was calibrated by simultaneously recording the sound using a calibrated measurement microphone so that the actual playback levels for each sequence are known. Computing the envelope of the signal (using the Hilbert envelope) and averaging 20 randomly generated pulses, as shown in Figure 2, shows how the method is able to capture the attack and release response of the device.

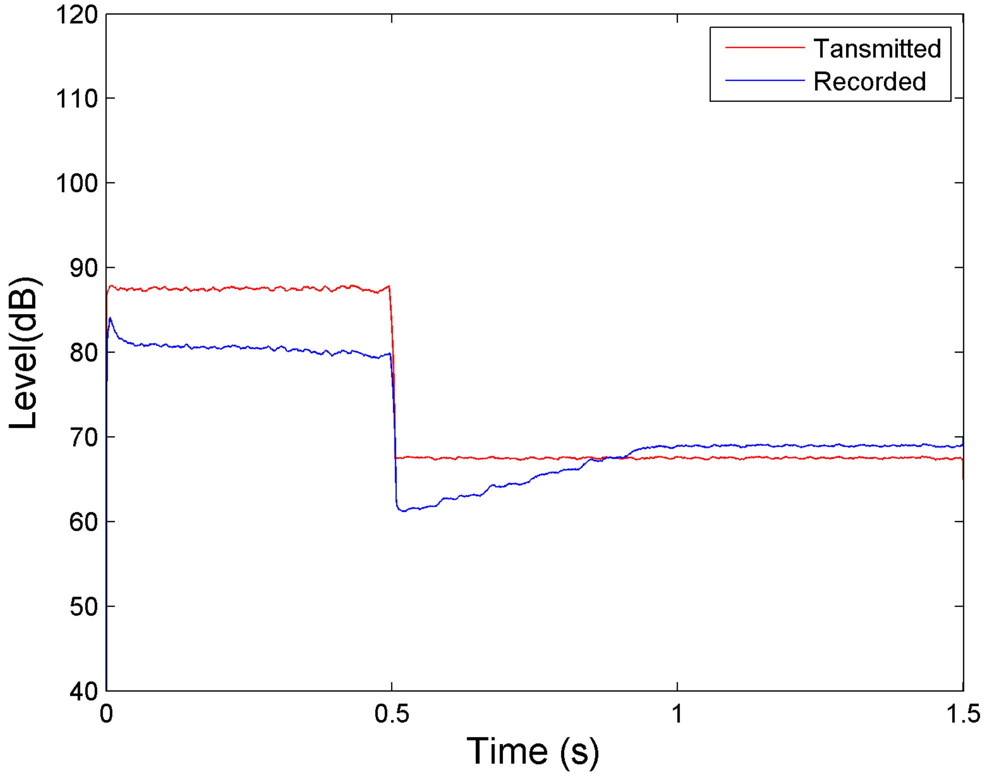


Figure 2. Average envelope of twenty pulses, showing both the device response in blue and the originally presented signal in red

As the threshold where the gain reduction or clipping occurs is not known, the sequence of 20 pulses was repeated at a number of different playback levels. Starting at the maximum level the loudspeaker was able to reproduce without significant distortion (Leq over the loudest part of the signal was 105 dB), and reducing the presentation level by 3 dB for each sequence. 15 presentations levels were applied. To time align all of the recordings a swept sine-wave pulse was sounded at the beginning of each measurement. The swept sine wave was played back at a relatively low level to try to ensure that the devices were not operating in a nonlinear region. A matched filter was then used to detect the time delay for each device and time align each test signal. The signal to noise ratio is increased by extending the length of the swept sine wave, this ensures a good estimate of the time delay. As the clocks of the various devices were not synchronised small amounts of relative drift between the transmitted and recorded signals were encountered. Figure 3 shows a time history of the 15 pulse sequences.

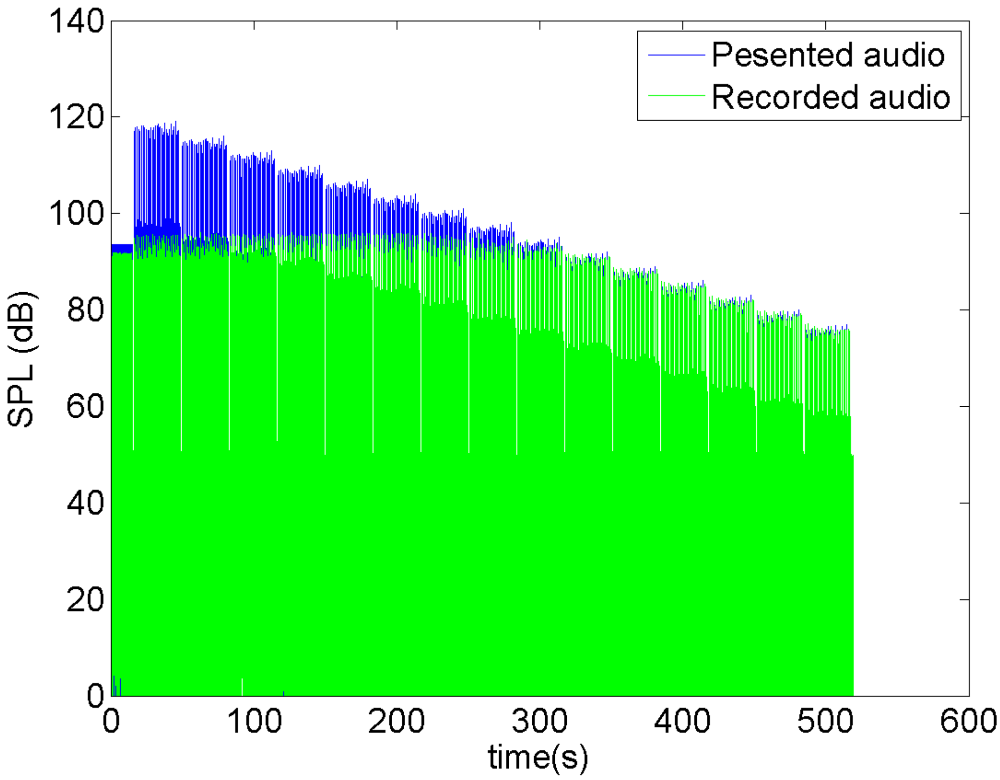


Figure 3. Complete signal time history of the transmitted and recorded signals

3.1. Estimation of static system parameters

Once the full noise pulse sequences were captured for each device, the waveforms were post-processed to quantify the non-linear dynamic characteristics. Firstly the signals were time aligned using the matched filter. Then each device's data was calibrated so that comparison between devices was more straightforward.

To calibrate each device the average level of the last half of each attack phase is computed for each playback level, the system is calibrated by finding a level when the device is operating in a linear mode. For each device this was selected by computing the gradient of the input-output gain curve: Selecting the playback level where the gradient is closest to one, selects calibration point where the system is operating in a linear fashion.

Figure 3 shows an example of a calibrated recording of one of the devices. The peak and rms levels in the last half of the attack phase averaged over all 20 pulses are shown in Figure 4. To compute the threshold the level where the device transitions between linear and non-linear operation or where the gain reduction is activated the error between the input and output level is used. The level above the identified linear region where the error between input and output level first exceeds 0.5 dB is selected as the threshold.

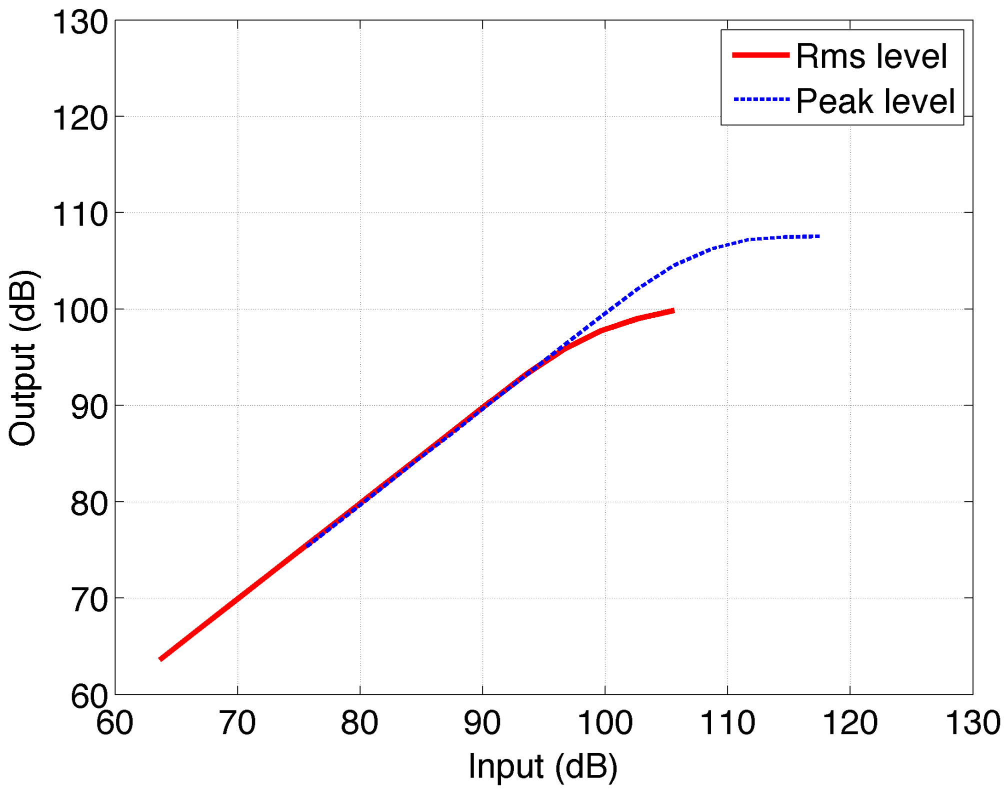


Figure 4. Peak and rms gain reduction curves

A simple gain curve model, with parameters including threshold, compression ratio and smoothing, is fitted to the computed gain curves.

3.2. Clipping detection

In the cases where the dynamic range control systems either are not present or the response is instantaneous; distortions such as soft or hard clipping can occur. A nonlinear transfer function, which maps one signal level to another, can be used to model the non-linearity. As synchronicity between the two signals can not be guaranteed, the amplitude histogram is used to calculate the transfer function using a model-based optimisation. A model of non-linearity was used where T is the linear threshold, R is the compression ratio, $x_v[n]$ is a vector of possible input levels.

$$y_i[n] = \begin{cases} x_v[n] & x_v[n] \leq T \\ T + \dfrac{x_v[n] - T}{R} & x_v[n] > T \end{cases}$$

$y_i[n]$ is the intermediate output level, a moving average filter was applied to smooth the gain curve, where the parameter M is the window size and controls the amount of smoothing.

$$y[n] = \frac{1}{M} \sum_{j=0}^{M-1} y_i[n+j]$$

The mapping of input to output level is then carried out by linearly interpolating $y[n]$.

This model has three parameters, threshold (T), compression ratio (R) and softness of the knee (M).

Model parameters are estimated by optimisation of a cost function. The cost function is the error between the output amplitude histogram and a simulated amplitude histogram. The search algorithm first performs a coarse search by evaluating the cost function over a parameter grid then a Simplex based optimisation using the coarse estimate as a starting point is performed. Figure 5 a) shows histograms of the input signal and the distorted signals. “Recorded” refers to the signal, which is captured by the device, “simulated” refers to the signal which is created by distorting the input signal by using the previously described model. Figure 5 shows an example of hard clipping where a sharp peak in the histogram is clearly visible due to the thresholding, Figure 5 b) shows resulting the optimised model’s gain curve.

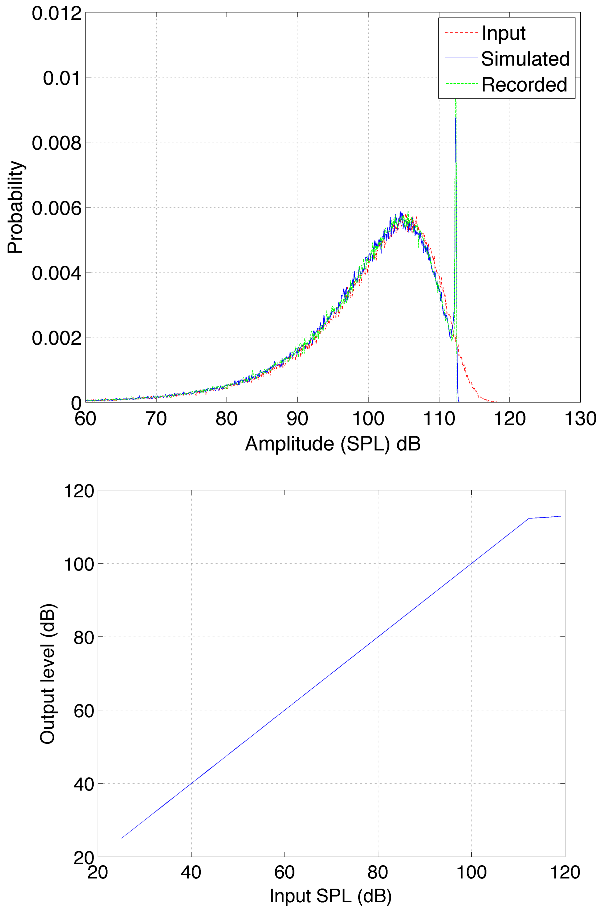


Figure 5. a) Amplitude histogram levels for the input, simulated and device recordings. B) non-linear model of clipping learnt from amplitude histogram

3.3. Temporal behaviour - attack

Figure 6 shows the complete dataset for a single device, the attack and release responses for all levels for one device. The dynamic attack and release responses are estimated by averaging the Hilbert amplitude envelope over the 20 examples. It is from this data that the parameters of a gain control system can be computed.

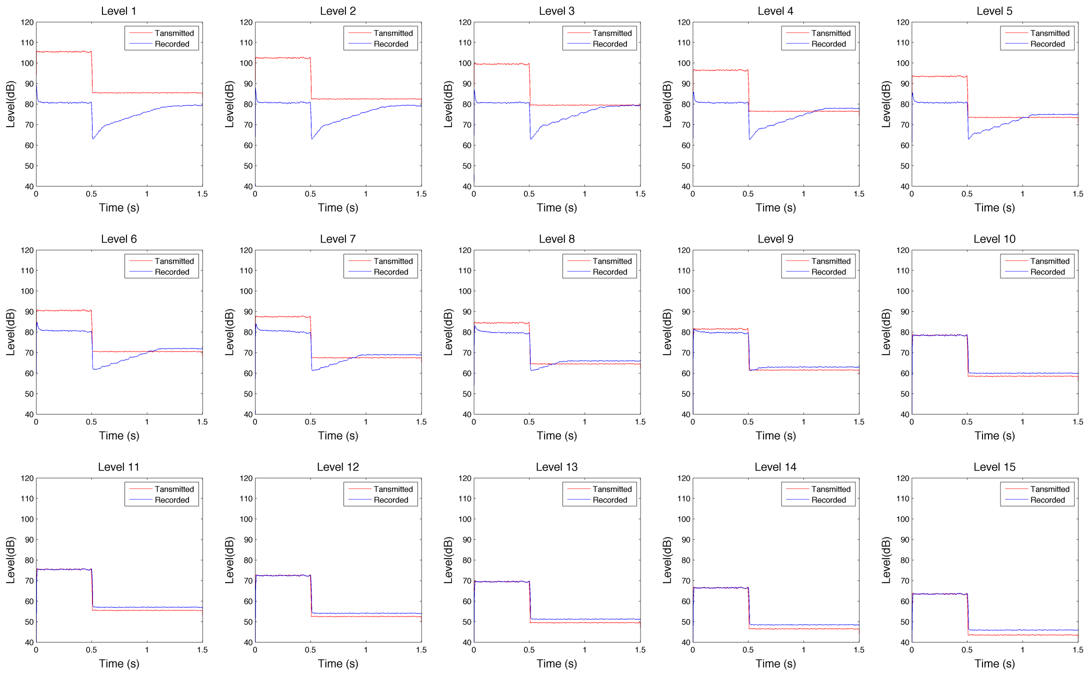


Figure 6. Results, all levels, both attack and release phases

To compute the attack curve the loudest pulse is used to ensure the threshold is crossed. The attack phase is isolated and windowed so that the maximum value in the first 0.1 s located as the start of the window. To compute the attack time a line is fitted to the last half of the pulse and extrapolated backwards. The attack time is taken as the time when the signal first comes within 1dB of the extrapolated line. To ensure that only significant events are identified attack curves with a dynamic range less than twice the standard deviation of the latter half of the attack phases are rejected. Figure 7 shows an example of the captured attack phase, where the attack time is 0.017s.

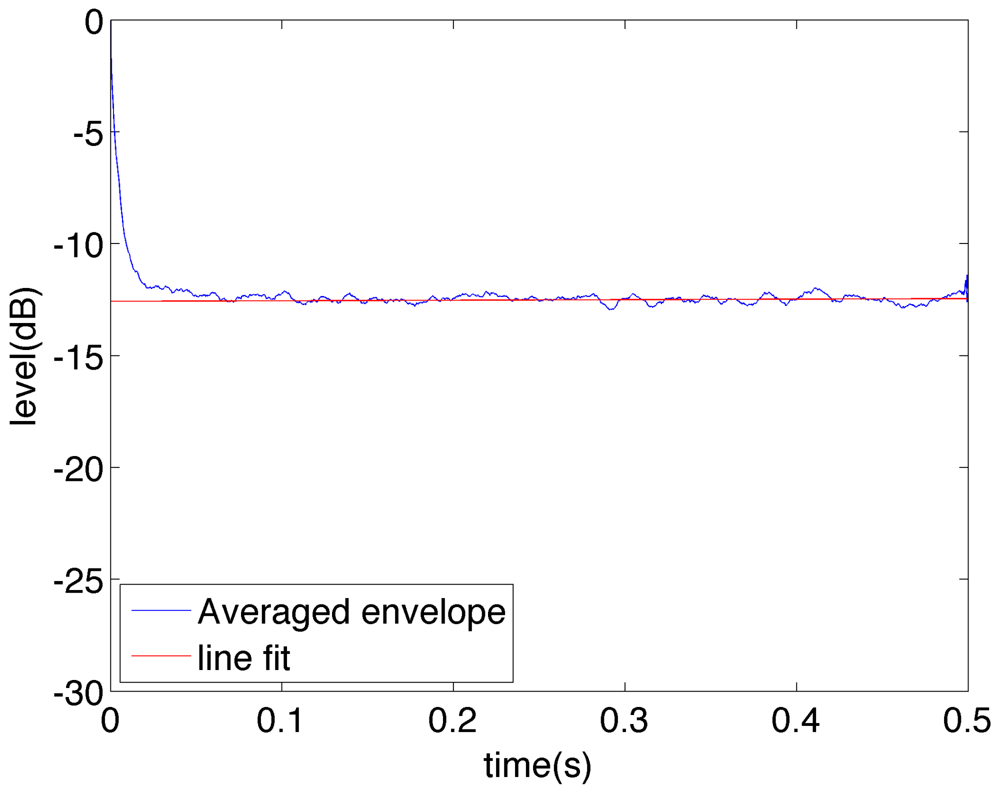


Figure 7. Attack phase, loudest level, the blue line shows the recorded signal, the red is a 1st order polynomial fitted to the latter half of the signals and is used to compute the attack time

3.4. Temporal behaviour - release

For the release phase, the peak threshold is used to indicate the playback level where the attack phase is above the threshold by around 10 dB. This ensures that the threshold has been crossed but also that during the release phase the level is under the threshold. The release phase fit is shown in Figure 8. This is computed in the same manner as the attack phase. The release time is 0.4 s.

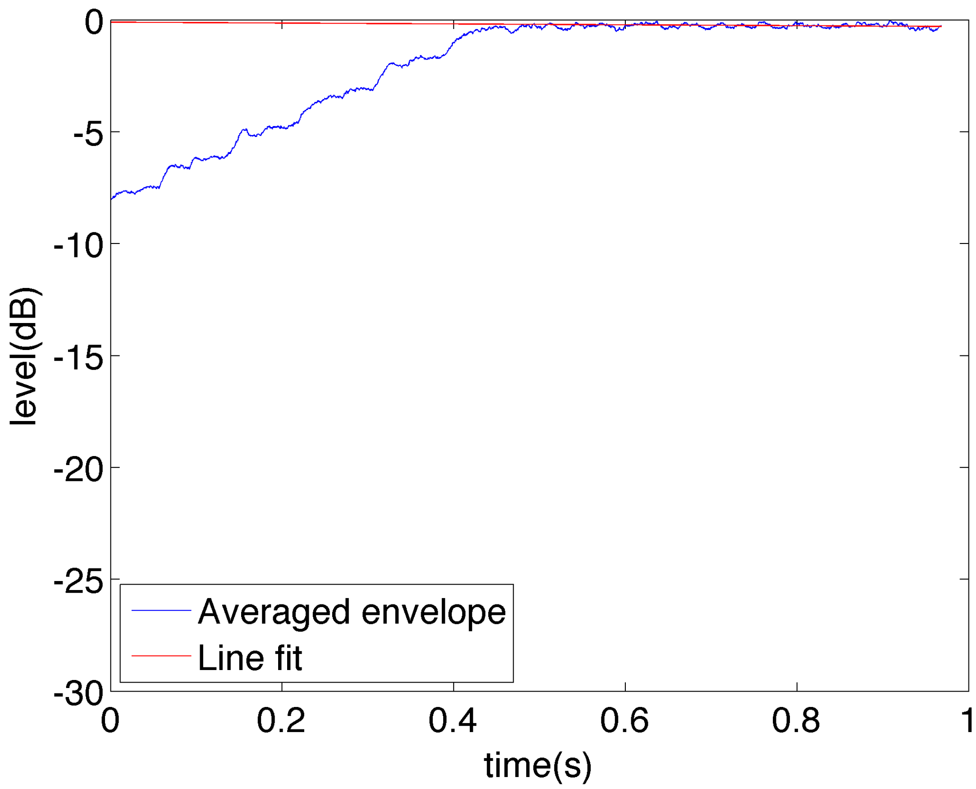


Figure 8. Captured release behaviour, the blue line shows the recorded signal, the red is a 1st order polynomial fitted to the latter half of the signals and is used to compute the release time

Automatic gain control systems will either respond to peak or RMS levels, in this report it was assumed to be RMS. (In future tests this could be examined using the method suggested previously, using a signal with constant peak level but variable RMS, in other words by quickly modulating the signal on and off. To achieve this the modulating frequency would have to be gradually increased, this would alter the RMS level while keeping the peak level constant).

4. Results

The following devices were tested, some with a number of settings.

Device	**Settings**
Cannon 550D	n/a
Edirol r44	Limiter off
Edirol r44	Limiter on
U87ai via Focusrite 2i4	n/a
SM57 via Focusrite 2i4	n/a
Zoom H2	No DRC
Zoom H2	AGC on
Zoom H2	Compressor on
Zoom H2	Limiter on
Zoom H4	No DRC
Zoom H4	AGC on

Zoom H4	Compressor on
Zoom H4	Limiter on
Google Nexus 4	AGC off
Google Nexus 4	AGC on
Sony camcorder	AGC off
Sony Camcorder	AGC on
Iphone	n/a

Table 1. Devices and setting used in experiment

After the recording was completed the audio was extracted from the devices and processed. The Appendix contains more details and figures, which show the temporal response, the static gain curve and the amplitude histograms. The following section summarises the response of each device.

4.1. Temporal behaviour, attack and release

The attack times of the tested devices are listed in Table 2. Items shown in bold represent attack and release times where the dynamic range of the behaviour is more than 1dB greater than the estimated stochastic variance. From this data the attack times, when present, vary betwen 5 to 17 ms. Release times vary from 30 to 400 ms. Some of the dynamic gain control systems did not show any detectable temporal response behaviour, this could be due to an extremely long attack time or even cases where the device permanently reduces the gain when overloaded.

		Attack time (s)	Attack dynamic range (dB)	Release Time(s)	Release dynamic range (dB)
Cannon 550D	n/a	**0.0171**	**12.51**	**0.3966**	**8.05**
Edirol r44	Limiter off	0.0001	2.43	0.0000	0.65
Edirol r44	Limiter on	0.0001	2.39	0.0000	0.62
U87ai	n/a	0.0005	2.99	0.0000	0.68
SM57	n/a	0.0001	2.82	0.0000	0.36
Zoom H2	No DRC	0.0001	2.42	0.0000	0.73
	AGC on	0.0002	3.29	0.0000	1.03
	Compressor on	**0.0014**	**3.97**	**0.3544**	**12.49**
	Limiter on	**0.0054**	**3.56**	0.0000	0.28
Zoom H4	No DRC	0.0001	2.66	0.0000	0.55
	AGC on	**0.0057**	**4.34**	0.0000	0.63
	Compressor on	**0.0063**	**10.98**	**0.1305**	**7.57**
	Limiter on	0.0003	2.83	0.0000	0.56
Google Nexus 4	AGC off	0.0001	3.30	0.0000	0.76
Google Nexus 4	AGC on	**0.0099**	**6.39**	0.0000	1.06
Sony camcorder	AGC off	0.0004	3.79	0.0000	0.35
Sony Camcorder	AGC on	**0.0081**	**5.74**	**0.0338**	**4.01**
Iphone	n/a	**0.0052**	**5.47**	**0.2226**	**3.76**

Table 2. Capture attack and release times, those marked in bold have a dynamic range greater than twice the natural standard deviation of the signal, and are deemed to be significant.

4.2. Static parameters, Gain, Threshold

Table 3 shows the threshold, compression ratio and smoothness of the gain curve for each device. The smoothness is the length of the window (M) used in the simulator, where 1 is no smoothing and higher numbers indicate increased smoothing. The devices marked in bold are those which have some temporal aspect to response, the other devices are better described by their non-linear behaviour which is the next table.

		Peak Thresh	Peak Compression Ratio	Smoothing	Rms Thresh	RMS compression ratio	Smoothing
Cannon 550D	n/a	**94.55**	**20.47**	**1**	**82.68**	**30.88**	**4.08**
Edirol r44	Limiter off	100.76	1930.90	1.00	96.32	3.37	1.00
Edirol r44	Limiter on	99.32	2587.20	1.00	94.87	3.31	1.41
U87ai	n/a	112.93	65.78	2	105.47	1049.30	1.00
SM57	n/a	112.09	62.59	1.00	105.36	1049.50	1.00
Zoom H2	No DRC	106.80	7.01	1.90	97.69	2.95	3.65
	AGC on	117.54	998.43	1.00	105.67	997.71	1.00
	Compressor on	**104.64**	**6.22**	**2**	**96.96**	**5.44**	**7.00**
	Limiter on	**108.08**	**26.37**	**3**	**98.73**	**12.45**	**2.14**
Zoom H4	No DRC	93.87	178.70	2	90.06	4.79	1.24
	AGC on	**118.00**	**1000.00**	**1**	**106.00**	**1000.00**	**1.00**
	Compressor on	**90.02**	**2.68**	**1**	**81.65**	**3.84**	**15.00**
	Limiter on	117.51	997.87	1.00	105.70	997.17	1.00
Google Nexus 4	AGC off	113.81	22.67	3.00	105.82	1049.60	1.00
Google Nexus 4	AGC on	**96.209**	**21465000**	**3**	**90.129**	**29.522**	**2.0253**
Sony camcorder	AGC off	111.07	5.0792	2	105.09	996.07	1.0074
Sony Camcorder	AGC on	**97.643**	**42.407**	**1**	**86.349**	**13.399**	**0.73278**
Iphone	n/a	**101.06**	**16.589**	**1**	**89.101**	**11.098**	**2.7799**

Table 3. Static gain performance, after attack phase and before release phase, bold items are records where the dynamic range processor has an attack or release time.

4.3. Static parameters, Clipping

The kurtosis of the waveform is a simple indication of the level of distortion. A Gaussian signal has a kurtosis of around 3. Examples, which deviate from this, are highlighted in bold; this indicates some form of non-linear distortion. A number of the devices show significant levels of hard clipping which are characterised by a sharp peak in the amplitude histogram (see Appendix), some indicate that there is some soft clipping, where the shape of the amplitude histogram is squashed into a smaller dynamic range. In some cases there is still some amplitude distortion even after the gain reduction is applied.

		Kurtosis	Peak Threshold	Compression ratio	Smoothing
Cannon 550D	n/a	3.05	92.60	1.79	27.56
Edirol r44	Limiter off	**1.24**	**100.72**	**342.37**	**47.00**
Edirol r44	Limiter on	**1.24**	**99.25**	**342.37**	**47.00**
U87ai	n/a	**2.65**	**112.86**	**17.01**	**34.71**
SM57	n/a	**2.51**	**111.92**	**9.66**	**46.90**

Zoom H2	No DRC	**2.22**	**100.13**	**1.42**	**38.70**
	AGC on	3.01	115.04	1.67	17.11
	Compressor on	**2.21**	**98.35**	**1.42**	**37.99**
	Limiter on	**2.29**	**98.87**	**1.40**	**35.87**
Zoom H4	No DRC	**1.19**	**93.90**	**239.38**	**1.00**
	AGC on	**3.28**	**116.18**	**2.80**	**17.03**
	Compressor on	3.02	99.39	1.75	13.71
	Limiter on	3.03	115.03	1.70	14.21
Google Nexus 4	AGC off	**2.67**	**113.40**	**9.23**	**81.61**
Google Nexus 4	AGC on	**1.56**	**92.21**	**6.91**	**54.31**
Sony camcorder	AGC off	**2.39**	**111.14**	**3.59**	**41.70**
Sony Camcorder	AGC on	2.90	96.29	1.73	30.63
Iphone	n/a	3.04	100.14	2.79	30.00

Table 4. Nonlinear behaviour of devices, bold indicates a deviation of the Kurtosis in the attack phase from Gaussian. This is an indication of some level of non-linear distortion being present, even after the gain reduction.

5. Conclusion

This report has detailed a method developed to quantify the range of responses that occur in portable audio devices when present with excessive sound pressure levels. Some devices exhibit forms of dynamic gain control, which act to reduce audible distortion by allowing gain to be adjusted gradually. The attack and release times of such systems range from 5 to 17 ms and 30 and 400 ms respectively. Some devices may also demonstrate a nonlinear gain curve with no attack or release but which try to limit audible distortion by using a compression ratio of between 1.4 and 10. While other systems have no protection and when presented with excessive sound levels will exhibit hard clipping.

6. Acknowledgements

We would like to thank EPSRC for funding the project (EP/J1013013/1)

7. References

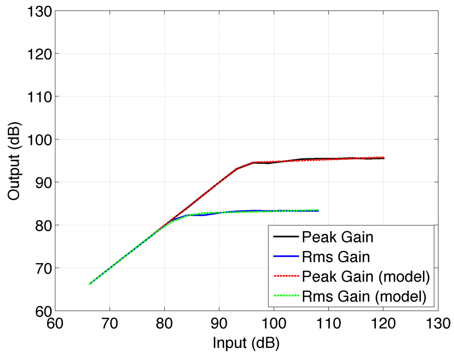
Bitzer, J., Schmidt, D., & Simmer, U. (2006). Parameter estimation of dynamic range compressors: Models, procedures and test signals. *Audio Engineering Society Convention 120*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=13653

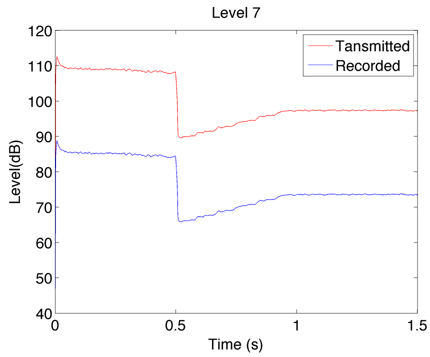
Cabot, R. (1987). Automated measurement of compressors and expanders. *Audio Engineering Society Convention 83*, *2513*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=4889

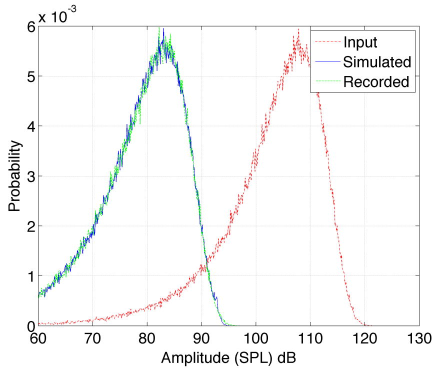
Schiffner, K., & Betz, T. (1996). *Measurement of Transient Responses in AGC Circuits using the Audio Analyzers UPL and UPD Products* (pp. 1–8).

8. Appendix

8.1. Cannon 550D

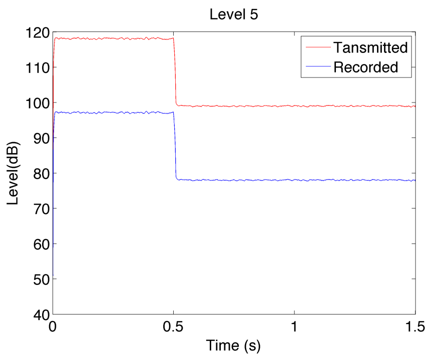


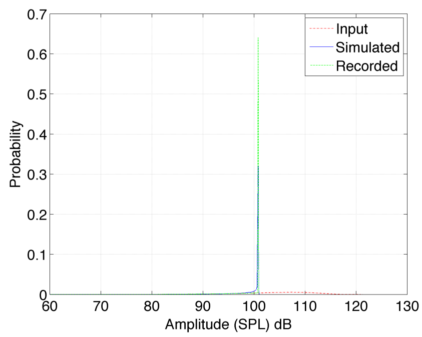


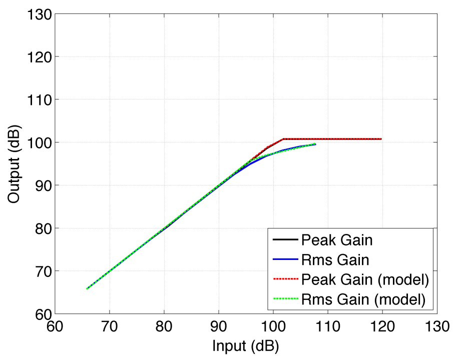


8.2. Edirol r44

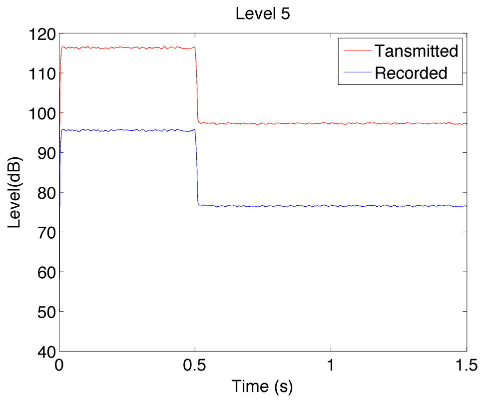
8.2.1. limiter off

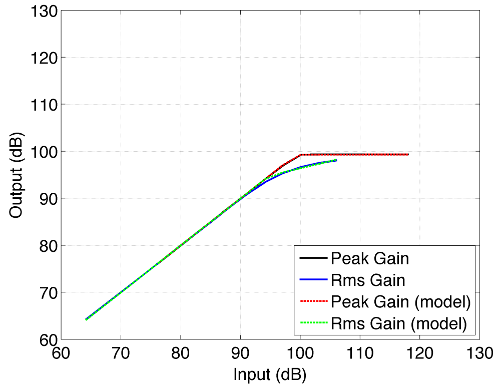


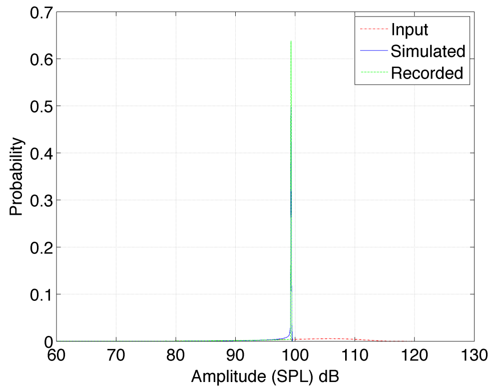




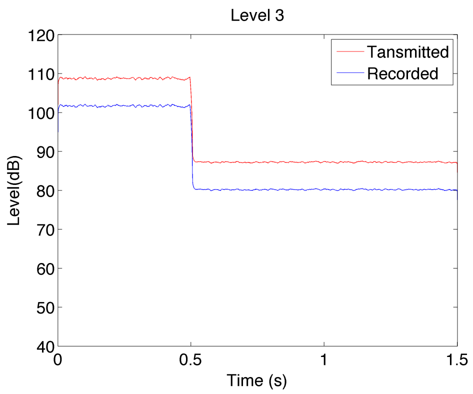
8.2.2. limiter on

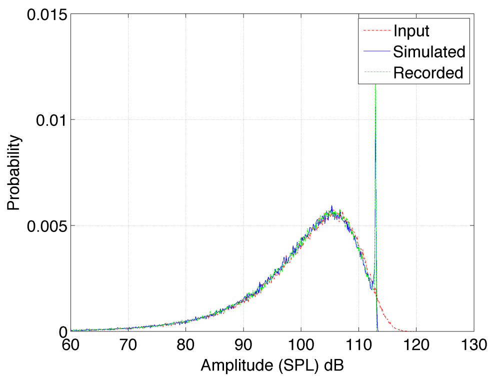


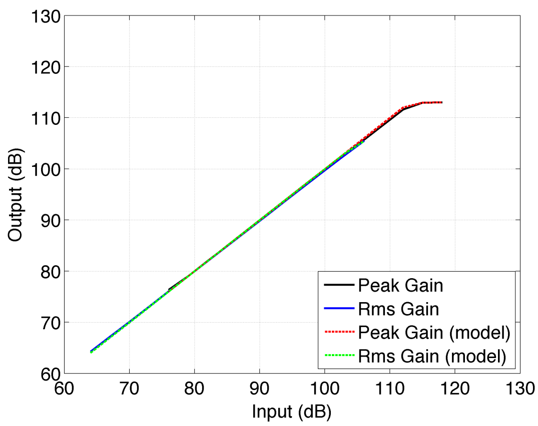




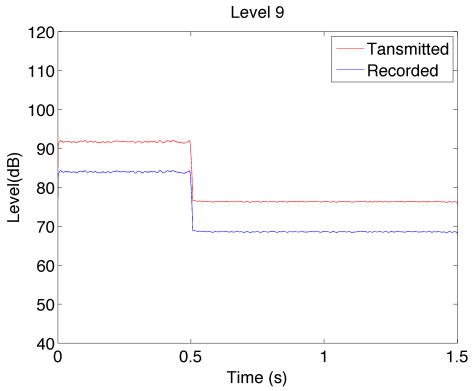
8.3. U87Ai



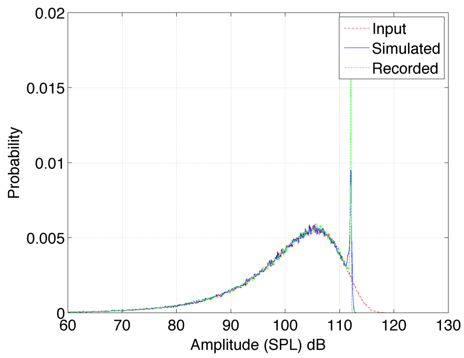




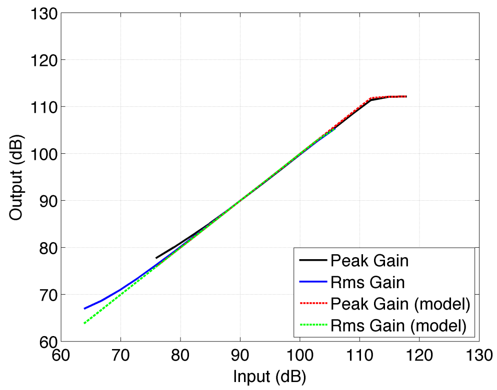
8.4. SM57



Level 9
Tansmitted
Recorded
Level(dB)
Time (s)



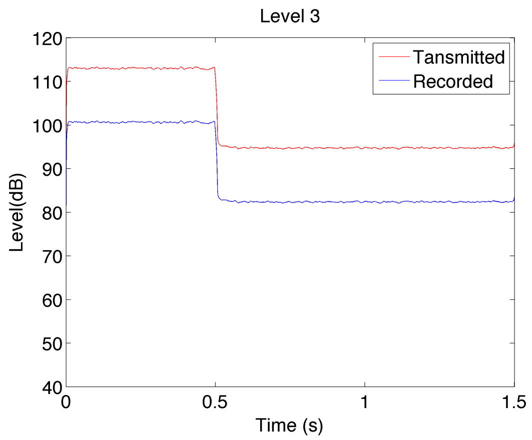
Input
Simulated
Recorded
Probability
Amplitude (SPL) dB

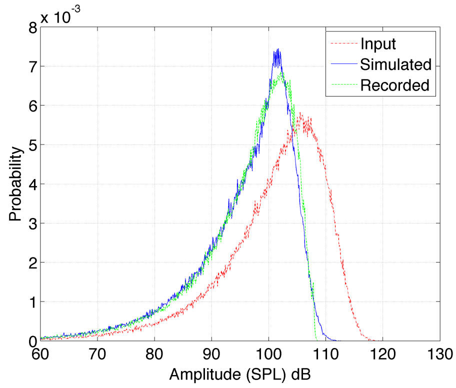


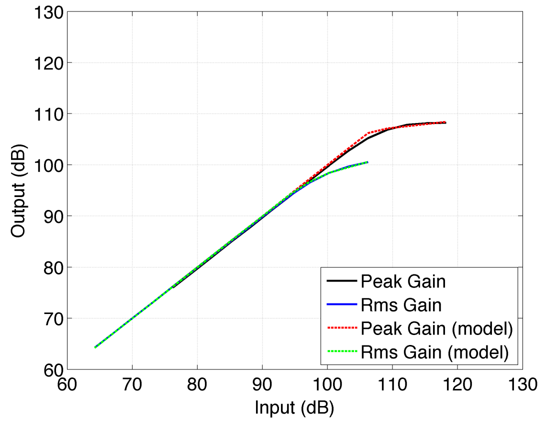
Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)
Output (dB)
Input (dB)

8.5. Zoom H2

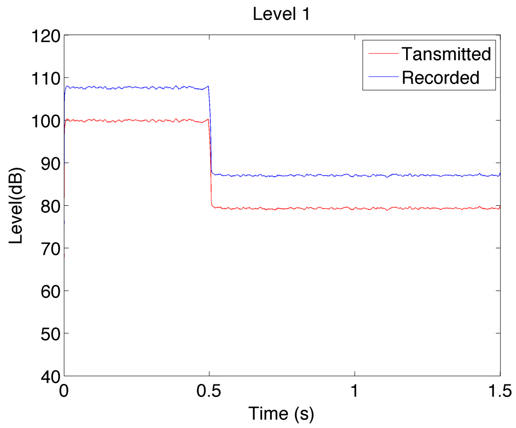
8.5.1. No dynamic processing

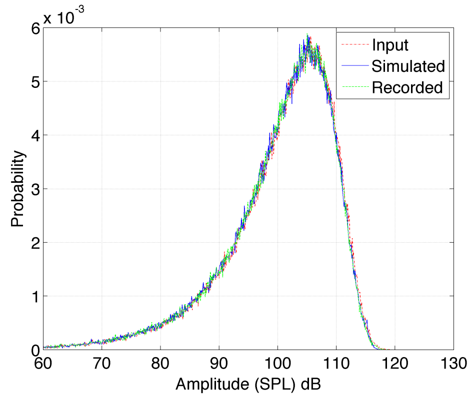


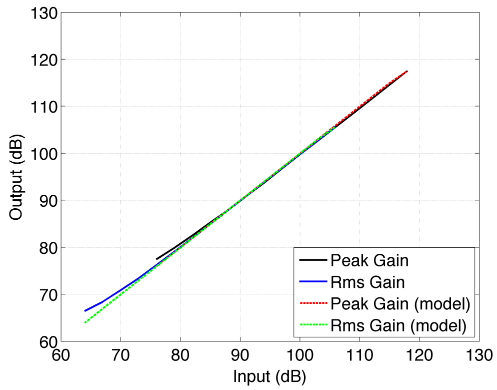




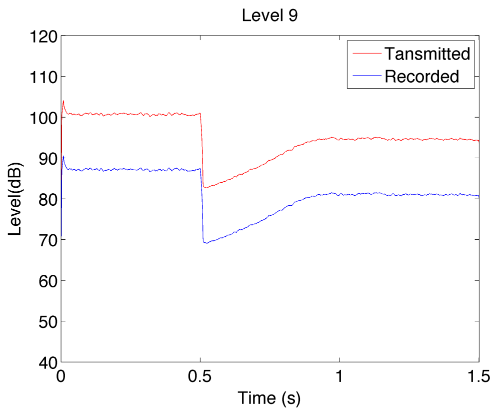
8.5.2. AGC on

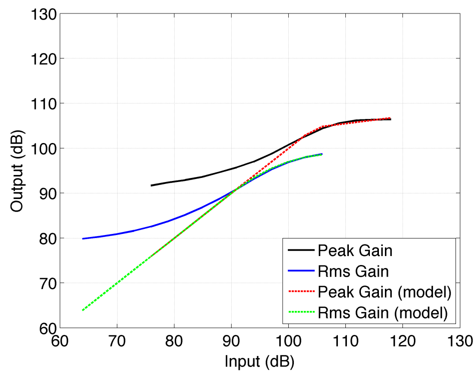


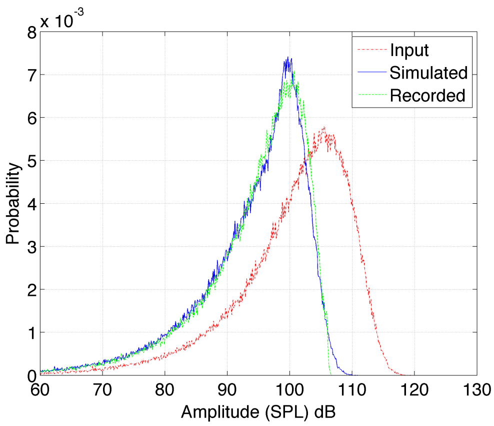




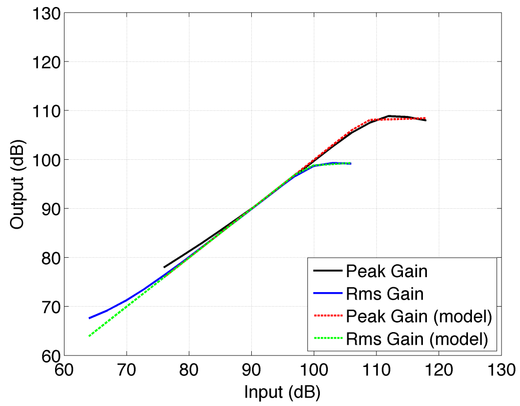
8.5.3. Compressor

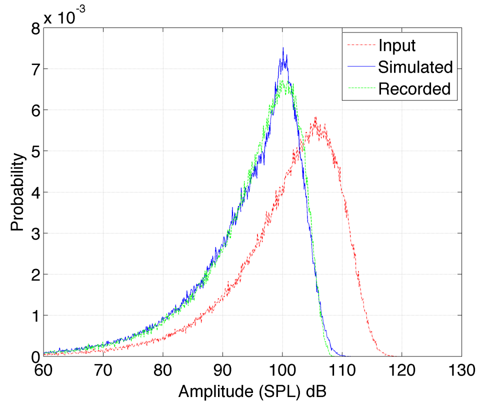


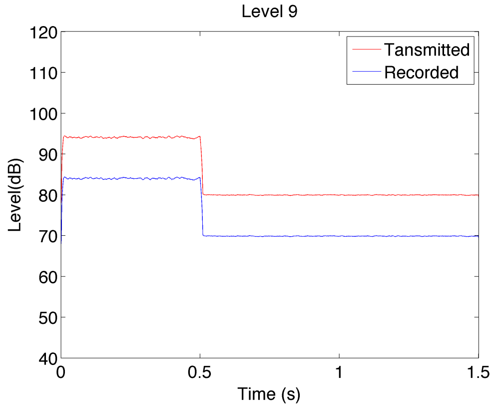




8.5.4. Limiter on

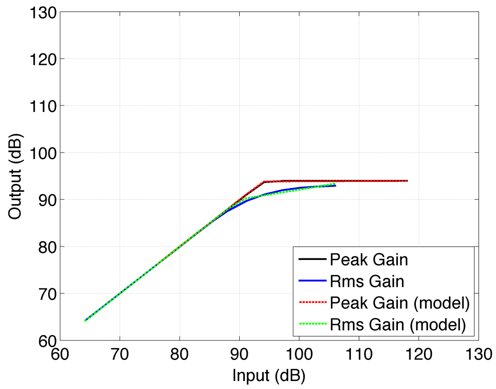


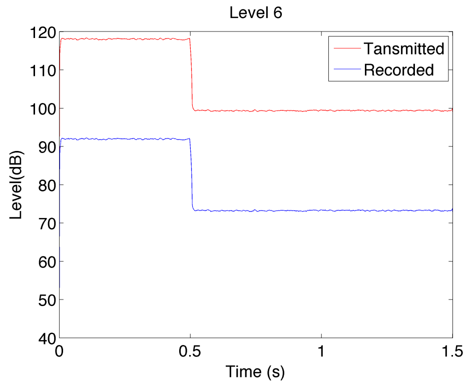


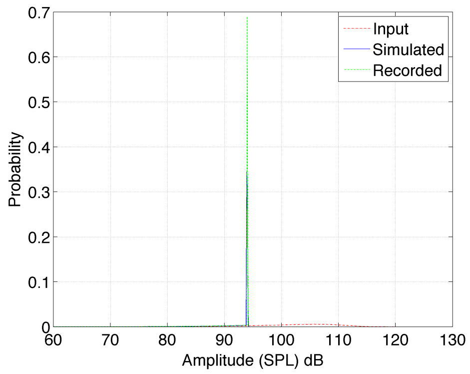


8.6. Zoom H4

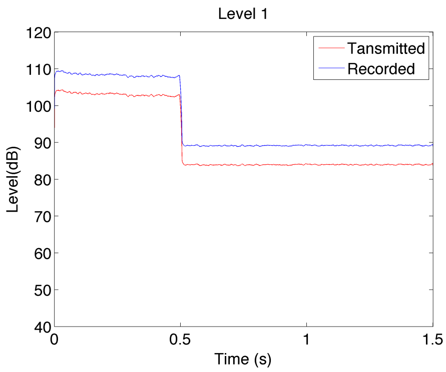
8.6.1. No dynamic processing

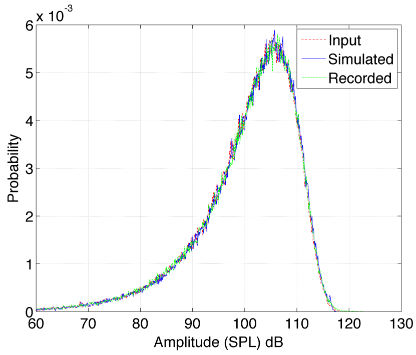

Output (dB)
Input (dB)
Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)
60
70
80
90
100
110
120
130

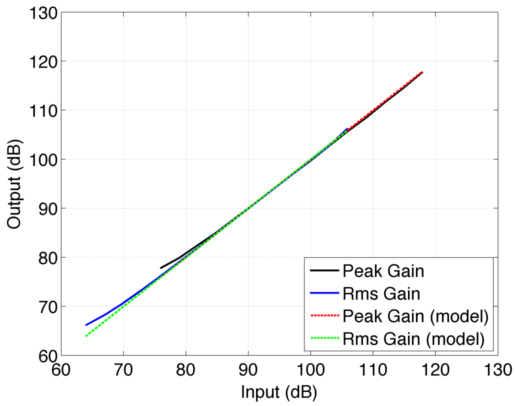

Level 6
Level(dB)
Time (s)
Tansmitted
Recorded
0
0.5
1
1.5
40
50
60
70
80
90
100
110
120


Probability
Amplitude (SPL) dB
Input
Simulated
Recorded
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
60
70
80
90
100
110
120
130

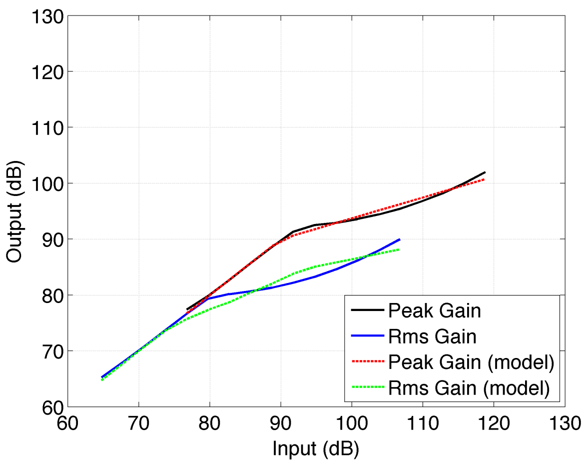
8.6.2. AGC on

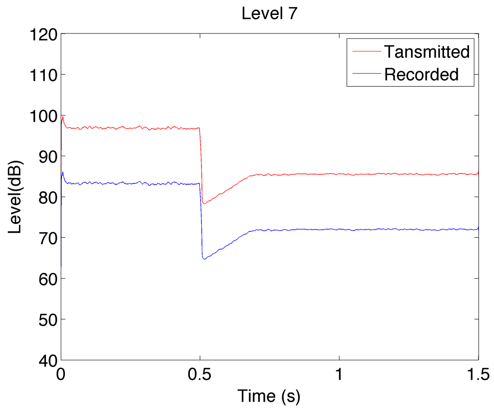


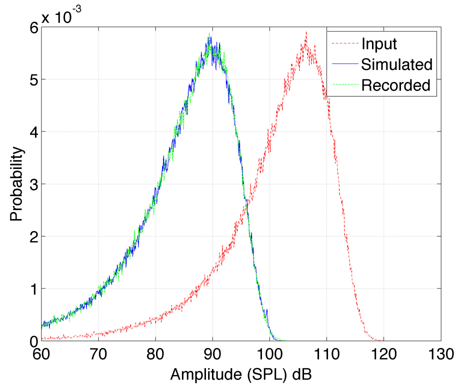




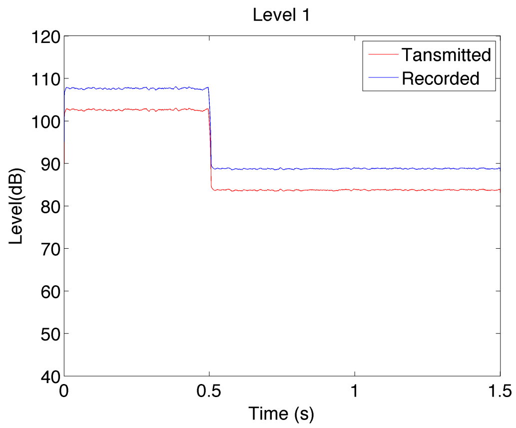
8.6.3. Compressor

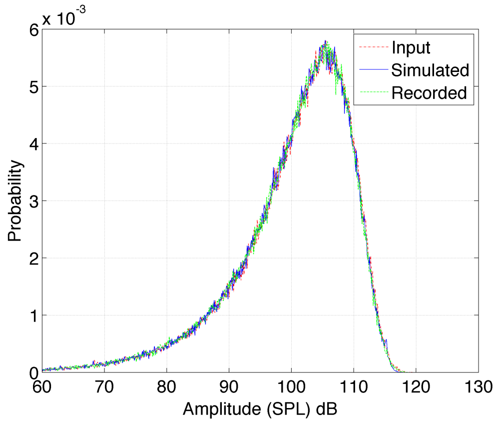


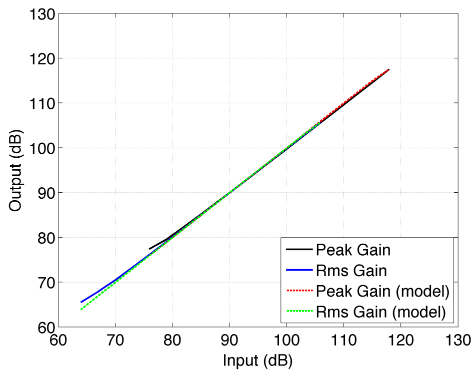




8.6.4. Limiter on

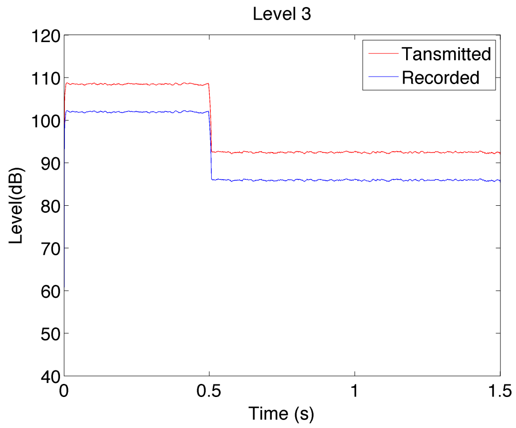


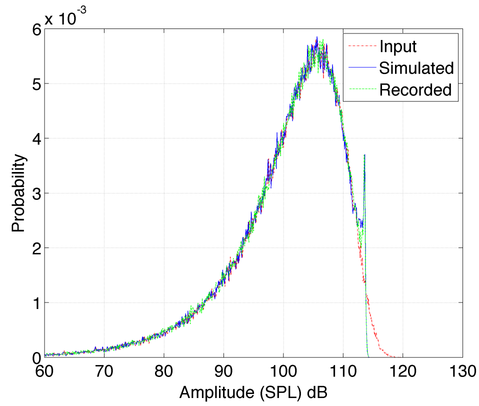


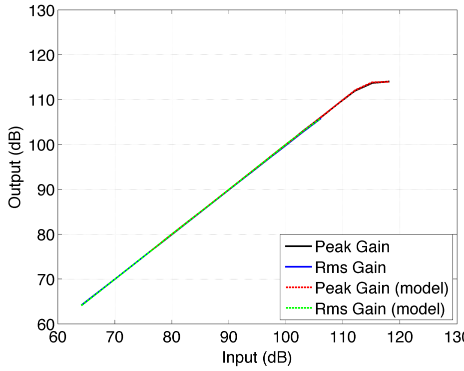


8.7. Google Nexus 4

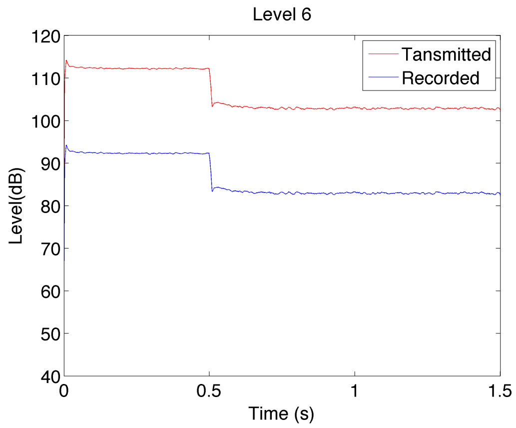
8.7.1. DRC off

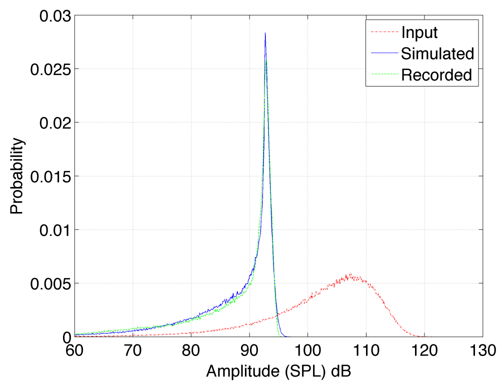

Level 3
120
110
100
90
80
70
60
50
40
Level(dB)
0
0.5
1
1.5
Time (s)
Tansmitted
Recorded

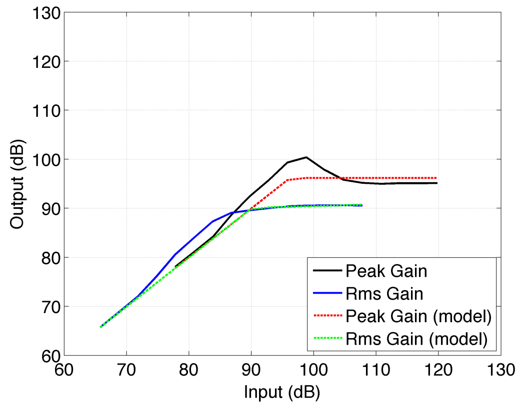

x 10^-3
6
5
4
3
2
1
0
Probability
60
70
80
90
100
110
120
130
Amplitude (SPL) dB
Input
Simulated
Recorded


130
120
110
100
90
80
70
60
Output (dB)
60
70
80
90
100
110
120
130
Input (dB)
Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)

8.7.2. DRC on

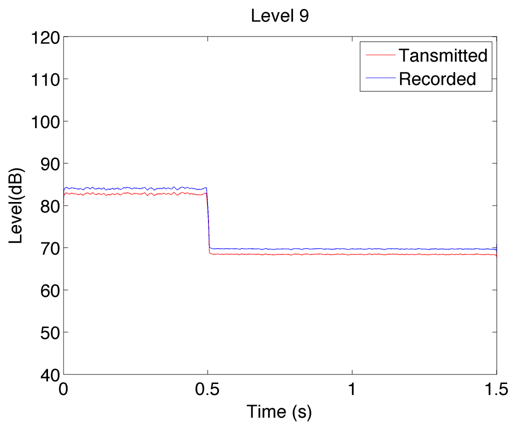

Level 6
120
110
100
90
80
70
60
50
40
Level(dB)
0
0.5
1
1.5
Time (s)
Tansmitted
Recorded

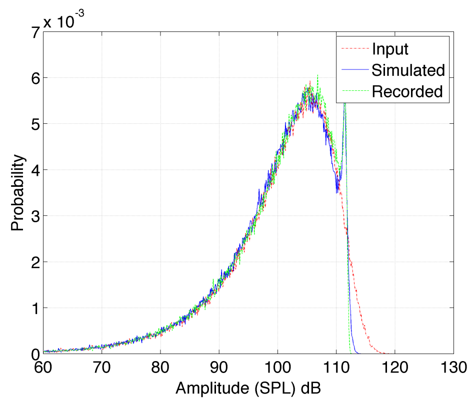

0.03
0.025
0.02
0.015
0.01
0.005
0
Probability
60
70
80
90
100
110
120
130
Amplitude (SPL) dB
Input
Simulated
Recorded

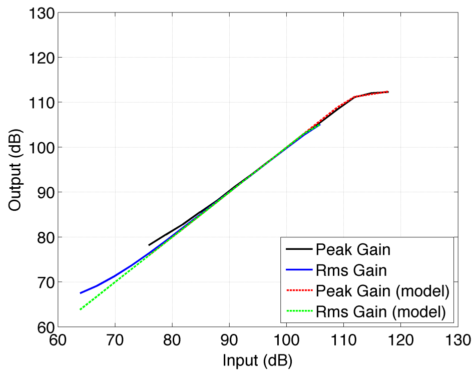

130
120
110
100
90
80
70
60
Output (dB)
60
70
80
90
100
110
120
130
Input (dB)
Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)

8.8. Sony Camcorder

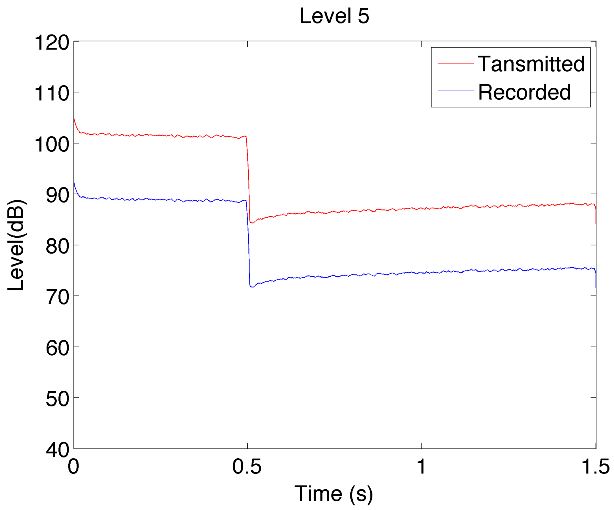
8.8.1. AGC off

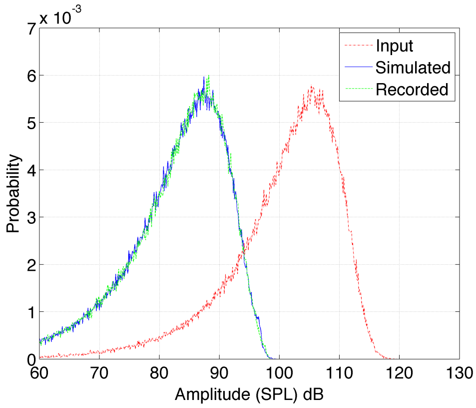


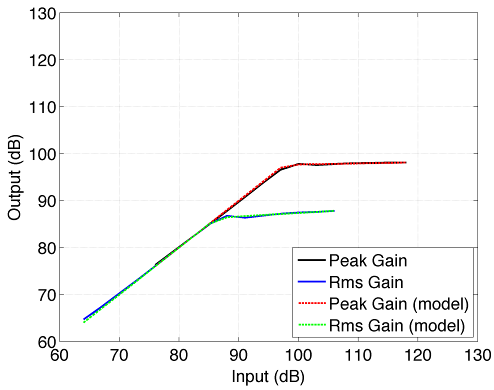




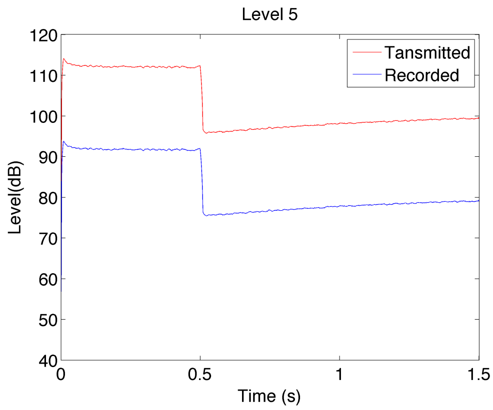
8.8.2. AGC on


Level 5
Tansmitted
Recorded
Level(dB)
Time (s)
120
110
100
90
80
70
60
50
40
0
0.5
1
1.5

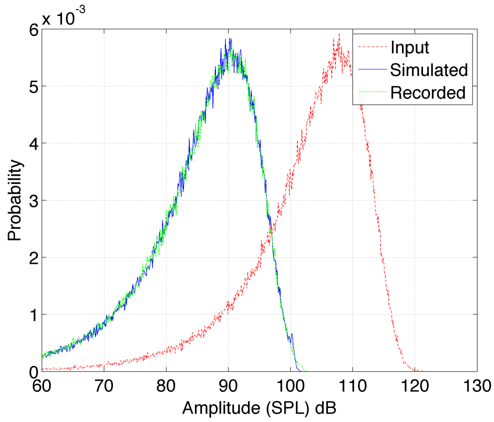

x 10^-3
Input
Simulated
Recorded
Probability
Amplitude (SPL) dB
7
6
5
4
3
2
1
0
60
70
80
90
100
110
120
130


Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)
Output (dB)
Input (dB)
130
120
110
100
90
80
70
60
60
70
80
90
100
110
120
130

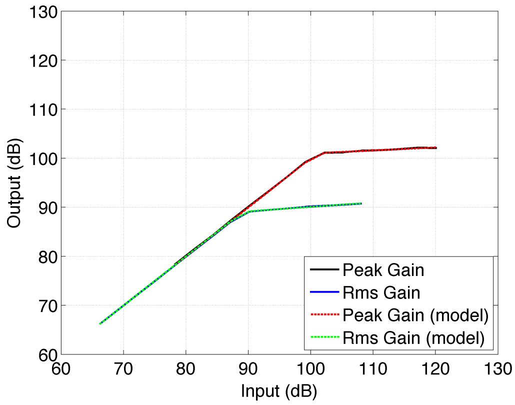
8.9. Iphone



Level 5
Tansmitted
Recorded
Level(dB)
Time (s)



x 10^-3
Input
Simulated
Recorded
Probability
Amplitude (SPL) dB



Peak Gain
Rms Gain
Peak Gain (model)
Rms Gain (model)
Output (dB)
Input (dB)