# OBJECTIVE EVALUATION OF THE AUDIBILITY OF TRANSIENT ERRORS IN AN ADAPTIVE A/D CONVERSION CHANNEL

*Niels Marker-Villumsen, Ivan H. H. Jørgensen and Erik Bruun*

Technical University of Denmark, Dept. of Electrical Engineering, DK-2800 Kgs. Lyngby, Denmark

nbvi@elektro.dtu.dk, ihhj@elektro.dtu.dk, eb@elektro.dtu.dk

## ABSTRACT

An adaptive analog-to-digital conversion channel for audio, using automatic gain control, generates transient errors that may be audible. Evaluating the audibility of such errors requires subjective evaluation using listening tests. From an electrical circuit design point-of-view this is not feasible, due to design time constraints.

This paper investigates the use of the model output variables (MOVs) from the Perceptual Evaluation of Audio Quality (PEAQ) method, for objectively evaluating the transient errors of the conversion channel, in order to optimize the design and reduce design time.

The objective method is compared with results from an alternative forced choice listening test. The comparison shows that the objective method can be used to evaluate the audibility of the transient errors; thus the method can be applied when designing the circuit implementing the channel.

***Index Terms—*** Objective Audio Evaluation, PEAQ, Alternative Forced Choice, Analog-Digital Conversion

## 1. INTRODUCTION

When designing electrical circuits, evaluating the circuits performance is crucial. For audio electronics, the circuits should be transparent from a signal quality point-of-view, to avoid reducing audio quality.

An adaptive analog-to-digital (A/D) conversion channel for audio has been developed, shown in Fig. 1. Based on [1] the main property of the channel is a dynamic range that is larger than the peak signal-to-noise ratio (SNR). This makes the current consumption in the analog part smaller than for a channel with a dynamic range equal to the peak SNR, since the current consumption is directly proportional to the peak SNR. However, as the channel is adaptively reconfigured, a transient error glitch is added to the output. This error may be audible, which is highly unwanted, and the audibility needs to be evaluated. Commonly the SNR and the total harmonic distortion (THD) are used as metrics when evaluating audio quality. Since they are only useful for steady state evaluation, carrying out listening tests is necessary. Unfortunately,

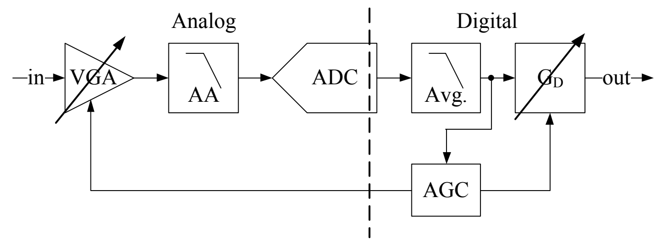



**Fig. 1**. Block diagram of adaptive A/D conversion channel

conducting a listening test is time-consuming, making trade-off evaluations in the design phase practically impossible. An objective evaluation using a computer based model would be preferred, for faster evaluation of the channel.

The mean squared error (MSE) would be a simple and useful metric to objectively evaluate the audibility of the transient error, by measuring the mean error energy. However, the calculated MSE can be the same for very different error signals [2], and does not take into account the frequency and temporal masking effects of the human hearing. To objectively evaluate the sound quality of high-fidelity audio systems with small impairments, the *Perceptual Evaluation of Audio Quality* (PEAQ) method can be used [3], [4]. Benjamin [5] used with good results the PEAQ method to evaluate audio quality degradation due to noise in analog-to-digital, digital-to-analog and sample-rate converters. Different parts of the human hearing are modelled and evaluated in the PEAQ method using model output variables (MOVs). Creusere and Hardin [6], [7] used the MOVs for objective evaluation of the audio quality of signals with temporally varying errors, also with good results.

Since the PEAQ evaluates systems with small impairments and grades the audio quality, it is not directly applicable for evaluating the audibility of the errors generated by the adaptive A/D conversion channel. This paper investigates the use of PEAQ MOVs to evaluate the audibility of transient errors generated by the adaptive A/D conversion channel. Specifically the *Maximum Filtered Probability of Detection* (MFPD) and *Average Distorted Block* (ADB) MOVs from the PEAQ method were used as they model the probability of detecting impairments present in the signal under test. To validate the usage of the MFPD and ADB, the computed results have been compared with the results of an alternative forced choice listening test.

## 2. ADAPTIVE A/D CONVERSION CHANNEL

An A/D conversion channel with a static gain is often used in e.g. microphones, which require a constant sound pressure level (SPL) sensitivity. For this type of channel, signal clipping occurs when the input signal becomes too large. This can be avoided by increasing the dynamic range of the channel, either by increasing the supply voltage, by decreasing the noise floor via increased bias currents, or by decreasing the overall channel gain. Unfortunately, these options may not be possible due to the specific application of the microphone.

Alternatively an adaptive A/D conversion channel can be used, with an analog and a digital part. A block diagram of this channel is shown in Fig. 1, consisting of an analog variable gain amplifier (VGA), an anti-aliasing (AA) filter, an analog-to-digital converter (ADC), an averaging filter, a digital gain block and an automatic gain controller (AGC). The overall gain of the channel is given as:

$$G_{tot} = G_a \cdot G_d \tag{1}$$

where $G_{tot}$ is the total channel gain, $G_a$ is the analog gain and $G_d$ is the digital gain. The AGC adjusts $G_a$ while simultaneously compensating by adjusting $G_d$, in order to achieve a constant channel gain. When the input signal level increases above a specific threshold level, the AGC decreases the analog gain and increases the digital gain, and vice versa when the signal level is reduced. In this manner the input dynamic range of the channel is increased, while maintaining a constant channel gain. The disadvantage is that the input referred noise of the ADC is increased when $G_a$ is reduced, causing an increase in the input referred noise of the channel for large input signals.

A more prominent problem is that when the channel gain is reconfigured, a transient error signal is generated due to the non-zero step response time from the output of the VGA to the input of the digital gain block. Assuming that the gain change occurs at $t = 0$, the error can be modelled as:

$$e(t) = \Delta G \cdot [h(t) - s(t)] \cdot x(t) \tag{2}$$

where $\Delta G$ is the change in gain, $h(t)$ is the Heaviside step function, $s(t)$ is the step response of the channel from the VGA output to the averaging filter output, and $x(t)$ is the input signal. This error equals a pulse, with a roll-off dependent on $s(t)$, and may be heard as a click. From (2), the peak value of the error is determined by $\Delta G$ and by $x(t)$. The value of $x(t)$ is related to the AGC threshold level, making both $\Delta G$ and $x(t)$ design parameters. The problem is to determine the optimum value for these parameters to avoid audible glitches in the output of the conversion channel. Thus an evaluation of the audibility of the transient errors is necessary.

## 3. OBJECTIVE EVALUATION OF ERROR SIGNAL

The PEAQ method evaluates the audio quality of a signal in several steps [3]. First the input signals (the reference and the signal under test) are transformed, using a model of the basilar membrane of the human ear, to generate excitation patterns. These are split into time-frames that are analysed in the frequency domain. The excitation patterns are further analysed for differences based on different aspects of the human hearing, represented using intermediate MOVs. The FFT based version of PEAQ uses two MOVs for modelling the probability of detection of a difference between the two signals: the MFPD and ADB.

For each frequency band in a frame, the probability of detecting the difference between the two signals is found, and used to determine the overall detection probability of the difference in each frame. The MFPD is calculated from the filtered probabilities as the maximum worst case filtered probability. The ADB models the distortion severity of the signal-under-tests as caused by the error signal. It is calculated as the average of the severity of distortion for each frame having a probability of detection above 50 %. For a more detailed description see [3], [8].

## 4. EXPERIMENTS

### 4.1. Listening Tests

To verify the results of the objective method when evaluating the transient errors of the A/D conversion channel, a subjective evaluation of the threshold level of hearing the error signal was carried out. A three interval, three alternative forced choice (3I3AFC) test was used together with the 1-up 1-down method [9]. The 1-up 1-down method was selected as it determines the point of 50 % probability of detection, as also used when computing the MFPD and ADB MOVs [4]. A 3I3AFC test was used instead of a 3I2AFC, to reduce the impact of random test answers on the overall test results.

The test subjects first did a training run, where feedback was given on the ability of the test subject to identify the correct error interval. The actual test consisted of three repeated runs for each of the three groups of test signals, each group using a different signal, resulting in a total of nine test runs. For all runs, the level of the transient error in the test signals was adjusted in steps of 4 dB and 2 dB during the search part, while 1 dB steps were used for the actual measurement part.

The scaling of the transient error signal was based on the method used in [10]. The generation of the tests signals is described further in Sec. 4.3.

The three test signal groups were created from 2.5 second long cut-outs of the *Double-bass*, *Tuba* and *English Male Speech* signals from the EBU Sound Quality Assessment Material CD [11]. The signals were selected in order to stress the conversion channel. The *Double-bass* and *Tuba* signals have low frequency content, which initial investigations showed decreases the error detection threshold level compared to signals with more high frequency content. The *English Male Speech* has a higher frequency content and is more complex, due to the many breaks and signal level variations. In this way the AGC would change gain settings more often, generating more transient errors in the output.

The test was carried out on a PC using the AFC MATLAB package [12]. A double-wall sound-attenuating listening booth was used for the test, and the signals were played back using a pair of *Sennheiser HD 580 Precision* headphones connected to a *RME DIGI96/8* 24 bit D/A converter with 48 kHz sample rate. The signal playback level was 68 dB SPL, with peak levels at 76 dB SPL, and the duration of the entire test was less than 1 hour for each test subject.

A total of 15 untrained test subjects were used, age range from 23 - 34, and all assumed to have normal hearing based on interviews prior to the tests. All experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-3-2013-004).

### 4.2. Model Simulations

For the objective evaluation of the test signals, the PEAQ implementation by Kabal [13] was used. This version implements the FFT based version of the PEAQ method, with the computed MOVs output scores directly available.

### 4.3. Generation of Test Signals

The test signals used for both the listening test and the objective method were generated using a high-level model of the A/D conversion channel. The VGA was modelled as a gain stage with a limiter function and gain settings from 0 dB to 18 dB in 6 dB steps. The AA filter was modelled as a 1st order low-pass filter with $f_{-3dB} = 200\ kHz$. The ADC was modelled as a linearized 4th order $\Delta\Sigma$ ADC, to only model the signal transfer function of the modulator. The averaging filter was implemented as a 16 tap FIR filter. Finally, the digital gain block was implemented as a multiplier with gain coefficients from 0 dB to 18 dB in 6 dB steps. The AGC was modelled with upper and lower signal threshold levels and with time-hysteresis to prevent the gain settings from constantly changing.

The transient errors for each input signal were found by subtracting a reference signal from the output signal of the A/D conversion channel model. Ideally the channel input signal would be used as the reference. However, due to the transfer function of the A/D conversion channel, the extracted error signal would also contain the difference caused by the phase shift of the channel. As only the transient error is of interest, the reference signal was generated using a reference model of the A/D conversion channel. The reference model was similar to the adaptive conversion channel, with the AGC and VGA limiter functions removed.

| | Test signal group | | |
|---|---|---|---|
| | *Double-Bass* | *Tuba* | *English-Speech* |
| $Q1$ | -34.9 dB | -35.6 dB | -21.1 dB |
| $Q2$ | -32.7 dB | -34.4 dB | -18.4 dB |
| $Q3$ | -30.8 dB | -33.1 dB | -15.9 dB |
| $\bar{x}$ | -32.7 dB | -34.1 dB | -18.7 dB |
| $s$ | 2.81 dB | 2.84 dB | 3.15 dB |

**Table 1**. Statistics for the transient error detection threshold levels from the results of the listening test

Both channel models were discrete time models implemented in MATLAB, using a sample rate of 2 MHz equal to the sample rate of the $\Delta\Sigma$ ADC. To simulate the A/D conversion channel using audio signals, the input signals were up-sampled from 48 kHz to 2 MHz. Using the extracted transient error signal, the test signals were generated by scaling the transient error signal from -60 dB to 6 dB and adding it to the reference signal. The model output signals were down-sampled to a 48 kHz sample rate and exported to WAVE files with 24 bit resolution.

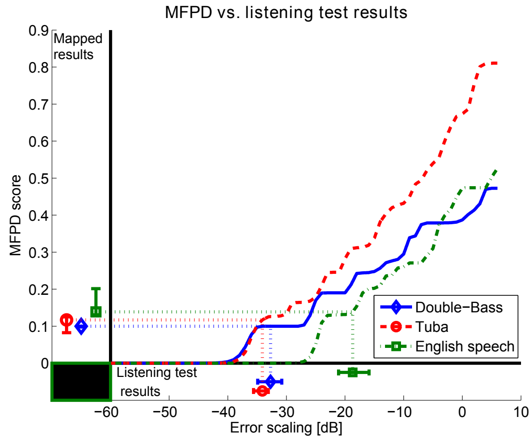
The peak value of the extracted transient error signals was approximately the same for all three input signals, as follows from (2) due to the fixed $\Delta G$ and AGC threshold levels. Nevertheless, the peak error value was not exactly the same for the three input signals since the AGC operates in discrete time. As a result, the error scaling rather than the peak error level was used for describing the transient error signal level in the test signals.

## 5. RESULTS

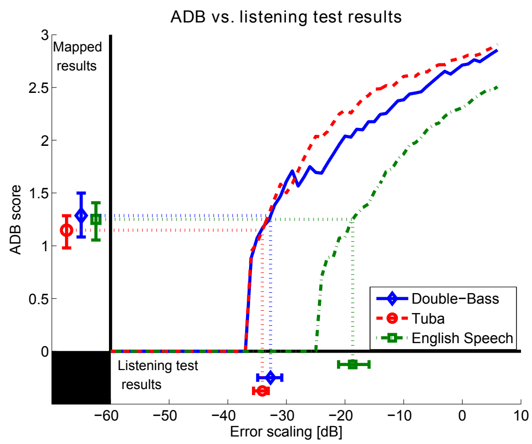
The mean detection threshold level for each test signal was determined based on the results of each test subject, and values for each test signal group were confirmed to be normal distributed by using normal probability plots. The mean threshold values are presented in Table 1, listing for each test signal group the 1st, 2nd and 3rd quartile ($Q1,\ Q2,\ Q3$), the sample mean $\bar{x}$, and the sample standard deviation $s$.

The calculated ADB and MFPD outputs for the three test signal groups are shown in Fig. 2. These outputs are the plotted curves, showing the equivalent MOV output value for each scaled error signal. The x-axis represents the error scaling factor used in the test signals, while the y-axis is the output value of the specific MOV.

The listening test results have been plotted in the bottom of the subfigures, and show for each signal group the found threshold mean and the 1st and 3rd quartiles. The mean value has been plotted with a marker while the 1st and 3rd quartiles are plotted as the edges of the variation lines. The plots on the left are the means and quartiles from the listening tests that have been mapped onto the MOV curves.




(a) MFPD




(b) ADB

**Fig. 2.** Calculated output values from MOVs

## 6. DISCUSSION

### 6.1. Bias Effects in Listening Tests

Due to the way the test signals were generated, the error was always present at the same time-instant for a given test signal group, independent of the error scaling. This effect is caused by the way the AGC works, as it triggers when the signal level crosses the specific threshold levels. Thus for a given input signal the error is always present at the same time-instant, independent of the size and shape of the error. Some test persons noticed this during the tests, resulting in a lower detection threshold. Consequently a lower mean threshold level was expected from the listening test in comparison to a test using non-repeated signals.

### 6.2. Evaluation of Results

The test results show that the error signal is easier to detect in the *Double-Bass* and *Tuba* signals compared to the *English Speech* signal; a consequence of signal frequency content and masking effects in the human hearing. Fig. 2 shows that for the ADB, the placement of the curves along the x-axis is similar to the test results. The MFPD output curves show the same trend, although not as clearly as for the ADB curves.

To compare the ability of the MOVs to evaluate the audibility of the transient error signals with results of the listening test, the test results were mapped onto the MOV output curves. The mappings showed that both MOVs generated output values in the same range for all test signals groups. In particular, the mapped variations for the ADB were closely matched, while the mapped results for MFPD had a larger variation. For the *Double-Bass* signal there was no variation in the MFPD mapping, because of the plateaus in the MFPD output curves. Generally these plateaus make the interpretation of the MFPD output difficult.

Based on these observations, we find that the ADB is an accurate method for objectively evaluating the transient errors of the adaptive A/D conversion channel, while the MFPD may be used for binary evaluation together with the ADB.

In relation to the optimization of the conversion channel in the design phase, it is relevant to consider which ADB and MFPD values one should target, to make the transient errors inaudible. The channel should be designed for the worst case situation, which from Fig. 2 is the threshold for the *Tuba* and *Double-Bass* signals. An option would be to aim for the lowest threshold levels found. However, during the tests the signals were played back at a specific SPL, which affected the error audibility. In contrast the MFPD and ADB output values were calculated from the unscaled test signals. It may be possible that the errors are audible if the signals are played back at a higher SPL; this makes it difficult to use the mapped MOV threshold values as a design target. Alternatively a goal is to achieve both ADB and MFPD outputs equal to zero; this equals a 50 % probability that there is no audible difference between the reference signal and the signal under test [4]. When designing and optimizing the adaptive A/D conversion channel this would be a conservative first design goal.

## 7. CONCLUSION AND FUTURE WORK

The adaptive A/D conversion channel has been evaluated using both subjective listening tests and objective computational methods. The results showed that the ADB is a good candidate for evaluating the audibility of the transient errors generated by the conversion channel while the MFPD is a less accurate tool. Based on the results, an output value of zero for both MOVs is a conservative first design goal when designing and evaluating the adaptive A/D conversion channel.

Future work includes evaluation of the objective method using more test signals, and also applying the method in an actual channel design and compare the computed results with results from a listening test.

## 8. REFERENCES


[1] M. Iwamatsu, "Analog-to-Digital Conversion Circuit," U.S. Patent 4 851 842, July 25, 1989.

[2] Z. Wang and A.C. Bovik, "Mean Squared Error: Love it or Leave it? A new look at Signal Fidelity Measures," *IEEE Signal Processing Mag*, vol. 1, pp. 98–117, 2009.

[3] T. Thiede et al., "PEAQ - The ITU Standard for Objective Measurement of Perceived Audio Quality," *J. Audio Eng. Soc.*, vol. 48, no. 1/2, pp. 3–29, 2000.

[4] "Method for Objective Measurements of Perceived Audio Quality," ITU-R BS.1387, 2001.

[5] E. Benjamin, "Evaluating Digital Audio Artifacts with PEAQ," in *Proc. 113th Audio Eng. Soc. Conv.*, Oct 2002.

[6] J.C. Hardin and C.D. Creusere, "Objective Analysis of Temporally Varying Audio Quality Metrics," in *Proc. 42nd Asilomar Conf.: Signals, Systems and Computers*, 2008, pp. 1245–1249.

[7] C.D. Creusere and J.C. Hardin, "Assessing the Quality of Audio Containing Temporally Varying Distortions," *IEEE Audio, Speech, Language Process.*, vol. 19-4, pp. 711–720, 2011.

[8] C. Colomes et al., "Perceptual Quality Assessment for Digital Audio: PEAQ-The New ITU Standard for Objective Measurement of the Perceived Audio Quality," in *Proc. 17th Int. Audio Eng. Soc. Conf.*, Aug 1999.

[9] H. Levitt, "Transformed Up-Down Methods in Psychoacoustics," *J. Acoust. Soc. Am.*, vol. 49, pp. 467–477, 1971.

[10] B. Feiten, "Measuring the Coding Margin of Perceptual Codecs with the Difference Signal," in *Proc. 102nd Audio Eng. Soc. Conv.*, Mar 1997.

[11] "Sound Quality Assessment Material Recordings for Subjective Tests," EBU - Tech 3253, 2008.

[12] S. Ewert, "AFC: The Psychophysical Measurement Package for MATLAB," http://medi.uni-oldenburg.de/afc/index.html.

[13] J. Kabal, "An Examination and Interpretation of ITU-R BS.1387: Perceptual Evaluation of Audio Quality," Tech. Rep., MMSP Lab, Dept. Elect. & Comput. Eng, McGill University, Canada, 2003.