ERROR RESILIENCE ENHANCEMENT FOR A ROBUST ADPCM AUDIO CODING SCHEME

Gediminas Simkus, Martin Holters, Udo Zölzer

Helmut Schmidt University — University of the Federal Armed Forces
Department of Signal Processing and Communications
Holstenhofweg 85, Hamburg, Germany

ABSTRACT

This work comprises an extension of a backward adaptive quantizer which is employed together with a robust lattice predictor in an ADPCM coding scheme. Predictors of the ADPCM audio coding schemes are often considered as the part most sensitive to transmission errors. Nevertheless, a single transmission error causes a short destabilization of the adaptive quantizer at the decoder side. Therefore, this destabilization boosts the deviation of the prediction filter at the decoder from the encoder side. Desired damping of the quantizer shortens synchronization periods. However, damping leads to degradation of the quantizer's adaptation properties and consequently to a decrease in audio quality. We show that the transmission of the quantizer's envelope to the decoder in short intervals and in combination with envelope error detector reduces the impairment of the reconstructed audio signal in noisy transmissions. An objective audio quality evaluation confirms significant quality improvement at BER higher than 10^{-4} and no quality degradation if an ideal channel is employed.

***Index Terms*—** Audio coding, error robust transmission, linear predictive coding, robust ADPCM

1. INTRODUCTION

Consider the following communication scenario. A digital wireless microphone is used by a stage performer. A wireless channel is employed for the transmission of encoded audio signal to decoder. At the receiver side the decoded signal is encoded again and returned to the performers in-ear monitor. If during audio encoding and decoding process an algorithmic latency is included then this latency should be less than 5 ms [1]. Furthermore, a wireless transmission channel has restricted channel capacity. Therefore, an audio coding scheme has to provide good quality for a desired bandwidth. Established lossy audio coding schemes such as MPEG-4 (AAC-LD) yield nearly transparent audio quality and guarantee high signal compression. However, the algorithmic latency equals to 20 ms [2]. This latency is present due to frequency transform based and block-wise processing. Therefore, frequency transform based audio coding schemes are latency critical for the previously named cascaded transmission scenario.

The wireless channel is inherently noisy. Besides bandwidth and latency constraints of the described real-time scenario, an audio coding scheme has to provide robustness against transmission errors. In Section 3 we propose an extension of the robust ADPCM coding scheme presented in [3]. The proposed coding scheme exploits redundancy which is added by an adaptive quantizer at the encoder side for better error resilient properties. This extension allows to improve audio quality in noisy transmissions and preserves good audio quality in noise-free channels.

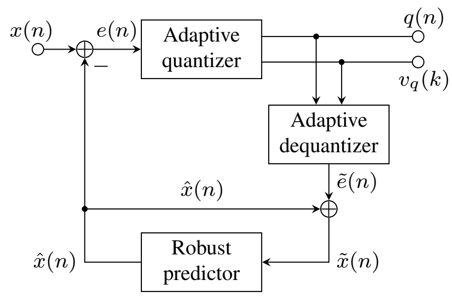
2. PRIOR LOW-LATENCY CODING SCHEMES

Low latency audio coding schemes developed over the past decades usually are based on linear predictive coding (LPC). LPC techniques are widely established in speech coding [4, 5, 6]. A typical application of a speech codec is mobile phone communication. Common speech coders such as G.726 [7] yield low latency, low bit rate and good speech quality. However, these speech coding schemes provide insufficient quality for other audio signals. Speech codecs usually add redundancy to guarantee robustness against transmission errors. Error correcting codes [8] are employed to reduce the impact of transmission errors. However, most methods introduce algorithmic delay to the audio coding scheme. Other speech codecs employ techniques [9] to detect so called click-noise in the decoded speech. The click-noise detector exploits slowly changing signal statistical properties. Therefore, click-noise can be concealed in a speech signal.

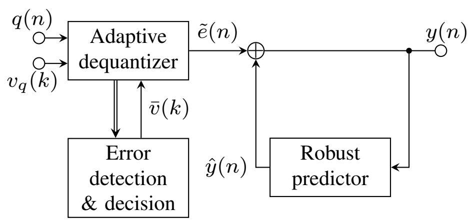
A perceptual audio coding concept [10] was achieved by employing pre- and post-filtering together with a base audio coding scheme. This block-based ultra low delay codec introduces an algorithmic delay of 6 ms at 32 kHz sampling rate. Techniques to ensure error robustness in noisy channels of that audio codec are discussed in [11].

A completely delay free audio coding scheme was proposed in [12]. A good audio quality has been attained by extending ADPCM by pre- and post-filters. However, this coding scheme presumes an error-free channel.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE



(a) Encoder



(b) Decoder

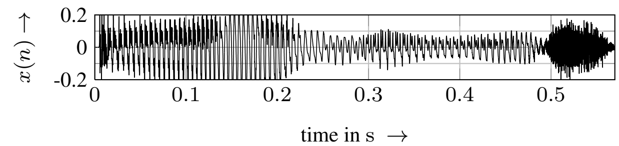
Fig. 1: Structure of the proposed coding scheme based on the robust ADPCM.

3. THE PROPOSED SYSTEM

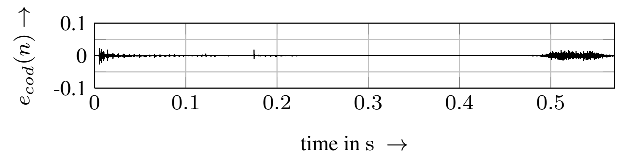
The proposed audio coding scheme depicted in Fig. 1 relies on error robust ADPCM [3]. Both, the prediction filter and the adaptive quantizer operate in backward manner. Therefore, there is no need to transmit side information in general. In the proposed coding scheme the quantizer scaling factor which is quantized $v_q(k)$ is transmitted to the decoder as side information in short intervals of r_v samples. Consequently, at every r_v-th sample the decoder is able to detect magnitude differences between received and calculated scaling factors. The error detection and decision is described in Section 3.2. The example in Fig. 2 justifies the effort of transmitting $v_q(k)$. The coding error $e_{cod}(n) = x(n) - y(n)$ of the robust ADPCM coding scheme in a error-free channel is depicted in Fig. 2b, where the original signal is shown in Fig. 2a. A single transmission error at the time 0.1 s alters the coding error as depicted in Fig. 2c. A transmission of $v_q(k)$ allows to stabilize adaptive quantizers at the encoder and decoder side. Therefore, the coding error of the proposed ADPCM coding scheme depends only on the predictor robustness properties. Hence, as shown in Fig. 2d, the corresponding coding error has a smaller power compared to $e_{cod}(n)$ in Fig. 2c.

3.1. Robust predictor

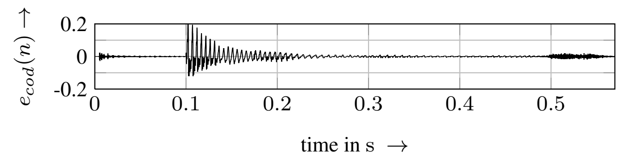
Now we give a brief description of the robust predictor which is employed in the proposed coding scheme (Fig. 1). The predictor is realized as a FIR filter in lattice structure [13, 14].



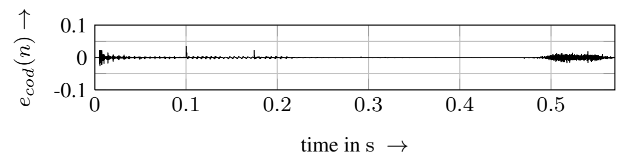
(a) Original signal



(b) Coding error in error-free transmission



(c) Coding error in noisy transmission based on robust ADPCM



(d) Coding error of noisy transmission based on proposed system

Fig. 2: The input signal word "animal" from the SQAM track 49 and corresponding coding error of the robust ADPCM. The corresponding coding error signals after the most-significant bit error at 0.1 s of robust ADPCM and proposed ADPCM.

The desired prediction is calculated by

$$\hat{x}(n) = \sum_{m=1}^{p} k_m(n) \cdot \alpha b_{m-1}(n-1), \tag{1}$$

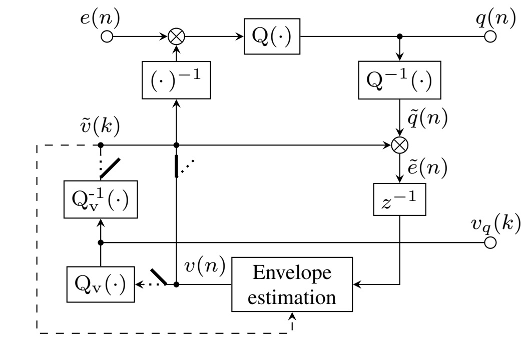
where $m = 0, \ldots, p$ denotes the lattice stage and p the desired prediction order. The reflection coefficients $k_m(n)$ are updated by the gradient adaptive lattice (GAL) [15] method iteratively by

$$\begin{aligned} k_m(n+1) = k_m(n) + \\ \mu_m(n) \cdot \big(f_m(n) \cdot \alpha\beta b_{m-1}(n-1) + b_m(n) \cdot f_{m-1}(n)\big). \end{aligned} \tag{2}$$

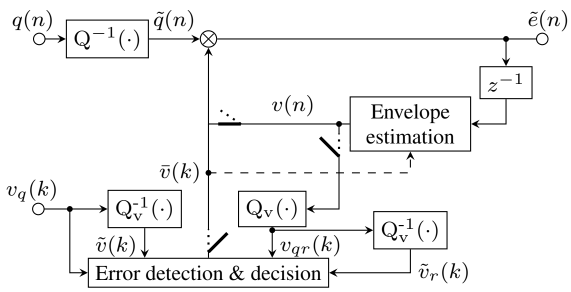
The forward $f_m(n)$ and backward $b_m(n)$ prediction errors are obtained by

$$f_m(n) = f_{m-1}(n) - k_m \cdot \alpha b_{m-1}(n-1) \tag{3}$$
$$b_m(n) = \alpha b_{m-1}(n-1) - k_m \cdot f_{m-1}(n). \tag{4}$$

The damping parameters $\beta < \alpha < 1$ ensure transmission error decay in the robust ADPCM coding scheme [3]. Appropriate coding parameters are $\alpha = 0.98$ and $\beta = 0.91$.



(a) Proposed adaptive quantization



(b) Proposed adaptive dequantization

Fig. 3: Structure of the backward adaptive quantizer. The dashed lines represent signal flow and switches position if current sample n is a multiple of r_v samples.

3.2. Proposed adaptive quantizer

The proposed backward adaptive quantizer is shown in Fig. 3. The solid lines represent the signal flow of conventional backward quantizer which is similar to the approaches presented in [16, 17]. The dashed lines illustrate adaptive quantizer operation when current sample n is a multiple of r_v, where r_v is a number of samples. This operation can be seen as synchronization of quantizers at encoder and decoder side at every r_v-th sample. The prediction error $e(n)$ is normalized at the encoder by its estimated envelope $v(n)$. At a synchronization point the envelope $v(n)$ is replaced by its quantized version $v_q(k)$ or rather $\tilde{v}(k)$. The envelope $v(n)$ is estimated by calculating the instantaneous power $v^2(n)$ of the reconstructed prediction error signal $\tilde{e}(n)$ as follows:

$$v^2(n) = \big(1-\lambda\big)\cdot v^{2\beta}(n-1) + \lambda\cdot\tilde{e}^2(n-1), \qquad (5)$$

$$\lambda = \begin{cases} \lambda_{AT} & \text{if } v^2(n-1) < \tilde{e}^2(n-1) \\ \lambda_{RT} & \text{otherwise} \end{cases} \quad \text{with } \lambda_{AT} > \lambda_{RT}.$$

The parameter λ controls how fast the estimate $v^2(n)$ follows the signal. The damping parameter β is chosen so that quantizers adaptability is not compromised. For every sample the prediction error $q(n) = \mathrm{Q}\left(\frac{e(n)}{v(n)}\right)$ and for every r_v-th sample the envelope $v_q(k) = \mathrm{Q}_\mathrm{v}\left(20\log_{10}(v(k))\right)$ are transmitted to the decoder. The payload quantizer $\mathrm{Q}(\cdot)$ is defined as in

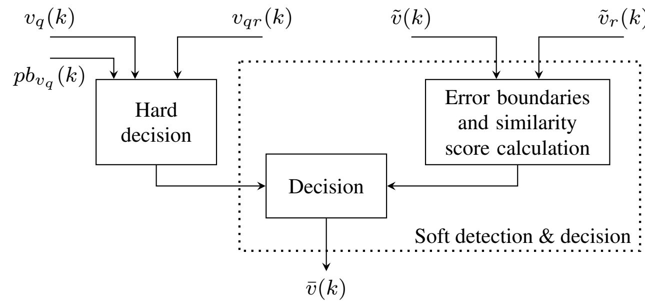


Fig. 4: Block structure of the transmission error detection and decision.

[16]. The overhead quantizer $\mathrm{Q}_\mathrm{v}(\cdot)$ is a logarithmically represented unipolar quantizer whose lower boundary is set to $20\log_{10}(v_{min})$, as $v(n)$ is bounded to $v(n) \geq v_{min}$.

The reconstruction of the received prediction error $q(n)$ at the decoder is done by $\tilde{e}(n) = \tilde{q}(n)\cdot v(n)$. The estimated envelope $v(n)$ is the same as at the encoder side, if previously no transmission error has occurred. According to Eq. 5, a transmission error at sample $\tilde{e}(n-1)$ causes an error at $v(n)$ and subsequent estimates and reconstructions. Therefore, as shown in Fig. 4 an error detection and decision (EDD) approach is employed. The EDD method accelerates the synchronization of adaptive quantizers at encoder and decoder side. Two synchronization points are separated by r_v samples. At the synchronization point the envelope $v(n)$ is estimated and $\tilde{v}(k) = \mathrm{Q}_\mathrm{v}^{-1}\left(v_q(k)\right)$ is received. An additional parity bit $pb_{v_q}(k)$ calculated from $v_q(k)$ is received from the encoder. The $pb_{v_q}(k)$ increases reliability of the received $v_q(k)$.

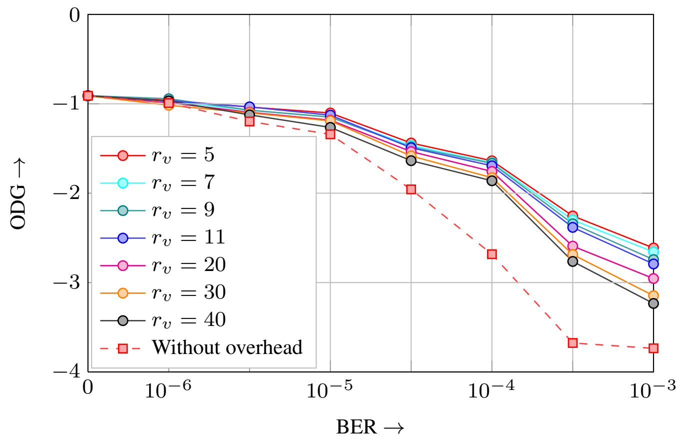
As depicted in Fig. 4 the EDD approach comprises hard error decision and soft error decision and detection. At the hard error decision side no error is found if the envelope quantized at the receiver side $v_{qr}(k)$ has identical word as the received word $v_q(k)$, where parity bit $pb_{v_q}(k)$ may be disturbed.

If the words $v_q(k)$ and $v_{qr}(k)$ are different, but the parity bit belongs to the received $v_q(k)$ then the decision for correctly received $v_q(k)$ can not be confirmed. In fact, the transmission errors may cause either even bit errors at $v_q(k)$ or bit errors at both $v_q(k)$ and $pb_{v_q}(k)$. Therefore, to avoid false positive hard error decision a so called similarity score and boundaries for the received envelope $\tilde{v}(k)$ are calculated. The similarity score gives a relative difference between $v_r(k)$ and $\tilde{v}(k)$ by

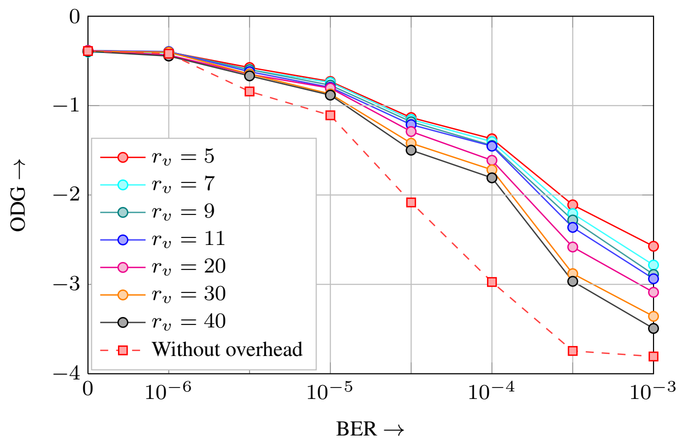
$$s_{score}(k) = \frac{|\tilde{v}_r(k) - \tilde{v}(k)|}{\frac{\tilde{v}_r(k)+\tilde{v}(k)}{2}}. \qquad (6)$$

The constraint $s_{score}(k) > s_{th}$ indicates to the erroneous envelope $\tilde{v}_r(k)$, where the threshold is set to $s_{th} = 0.1$.

As the envelope estimator is able to follow the signal $\tilde{e}(n)$ only at a certain range, possible boundaries for a received $\tilde{v}(k)$ are estimated similarly as in Eq. 5. Finally, the decision for the current $\bar{v}(k)$ is done based on collected hard er-



(a) Payload word length 3 bit



(b) Payload word length 4 bit

Fig. 5: Mean objective evaluation results at different BERs. The proposed ADPCM with overhead transmission at every r_v-th sample and the robust ADPCM without use of any additional side information are evaluated.

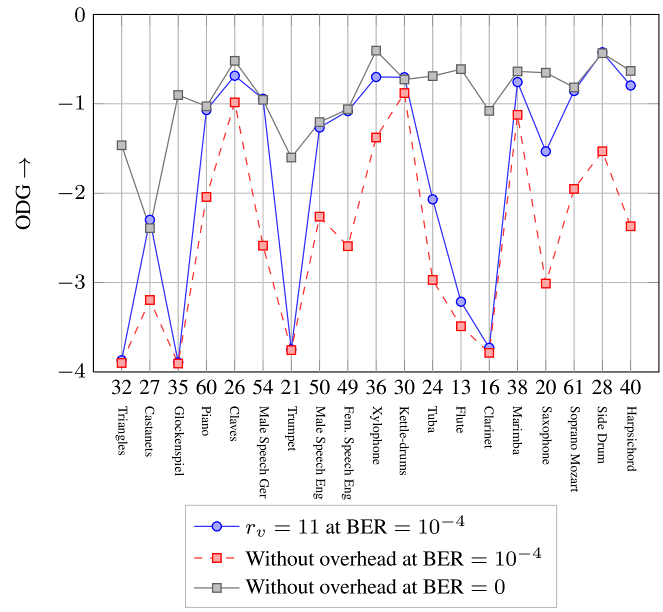


Fig. 6: ODG results of selected SQAM tracks coded with the proposed approach and the robust ADPCM at BER $= 10^{-4}$ and employing error-free channel. Payload word length of 3 bit and overhead of 0.63 bit/sample.

ror decision, calculated $s_{score}(k)$ and envelope boundaries. Therefore, if an error is detected at $v_q(k)$, then $\bar{v}(k) = \tilde{v}_r(k)$. If an error is detected at $\tilde{v}_r(k)$, then $\bar{v}(k) = \tilde{v}(k)$, else $\bar{v}(k)$ is set to the last known correctly received $\tilde{v}(k)$.

4. EVALUATION RESULTS

The perceptual audio quality is evaluated by the ITU-R BS.1387-1 (PEAQ) method [18] employing error-free and erroneous transmission channels. The objective difference grade (ODG) on scale from -4 (very annoying impairment) to 0 (imperceptible impairment) gives the perceptual audio quality measure of a test signal. All tracks from SQAM CD [19] are coded by the robust ADPCM and the proposed ADPCM coding scheme. The monaural audio excerpts (starting from 0.5 s and 10 s long) with sampling frequency 44.1 kHz are coded using payload of 3 and 4 bit/sample. Side information $v_q(n)$ is coded at r_v-th sample with the word length of 6 bit and additional parity bit. The robust and proposed ADPCM schemes parameters are set to the values as in [3] and [16].

The audio quality evaluation results at different BERs of the proposed and the robust ADPCM coding schemes are shown in Fig. 5. The proposed method starting from BER of $5 \cdot 10^{-6}$ achieves slightly better audio quality in comparison to the robust ADPCM codec. The audio quality of signals coded by the robust ADPCM at BER of 10^{-4} decreases rapidly. However, the results of the proposed approach at diverse r_v stay significantly higher. As shown in Fig. 6, the proposed method is incapable to achieve ODG improvement for several slowly decaying tonal signals such as Tracks 32, 35, 21 and 16. Nevertheless, some signals yield similar ODG at noisy and noise-free channels. Additionally, selected audio examples are available for listening on the website [20].

5. CONCLUSIONS

The proposed coding scheme compared to the robust ADPCM achieves significantly better audio quality in noisy channels. The proposed system adds an overhead which depends on the chosen transmission rate. If the overhead is transmitted in parallel to the payload, then the proposed approach may be seen as latency-free codec. To sum up, the proposed ADPCM achieves in average not annoying to slightly annoying impairment of the decoded signal up to BER of 10^{-4} if the payload is coded with a word length of 3 or 4 bit/sample.

6. REFERENCES

[1] Aki Härmä and Unto K. Laine, "Warped low-delay CELP for wideband audio coding," in *Audio Engineering Society Conference: 17th International Conference: High-Quality Audio Coding*, Aug. 1999.

[2] Gayer Marc, Lutzky Manfred, Schuller Gerald, Krämer Ulrich, and Wabnik Stefan, "A guideline to audio codec delay," in *Audio Engineering Society Convention 116*, May 2004.

[3] Gediminas Simkus, Martin Holters, and Udo Zölzer, "Error robust delay-free lossy audio coding based on ADPCM," in *Proc. of the 16th Int. Conference on Digital Audio Effects (DAFx-13)*, Sep. 2013.

[4] P. Elias, "Predictive coding–I," Mar. 1955, vol. 1, pp. 16–24.

[5] B. S. Atal and M. R. Schroeder, "Adaptive predictive coding of speach signals," 1970, The Bell System Technical Journal.

[6] Jerry D. Gibson, "Sequentially adaptive backward prediction in ADPCM speech coders," Jan. 1978, vol. 26, pp. 145–150.

[7] International Telecommunication Union, "40, 32, 24, 13 kbit/s adaptive differential pulse code modulation ADPCM. ITU-T recommendation G.726," *General Aspects of Digital Transmission Systems, ITU-T*, 1990.

[8] R. Kohno, S. Pasupathy, H. Imai, and M. Hatori, "A robust ADPCM system using an error-correcting code," in *Acoustics, Speech and Signal Processing. ICASSP 1986. IEEE International Conference on*, 1986, vol. 11, pp. 3091–3094.

[9] S. Kubota, A. Dobashi, M. Suzuki, T. Hasumi, and S. Kato, "Improved adpcm voice signal transmission employing click-noise detection scheme for tdma-tdd personal communication systems," *Vehicular Technology, IEEE Transactions on*, vol. 46, no. 1, pp. 108–113, 1997.

[10] Gerald Schuller and Aki Hanna, "Low delay audio compression using predictive coding," in *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, May 2002, vol. 2, pp. II–1853–II–1856.

[11] S. Wabnik, G. Schuller, and F. Kraemer, "An error robust ultra low delay audio coder using an ma prediction model," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, 2009, pp. 5–8.

[12] Martin Holters and Udo Zölzer, "Delay-free lossy audio coding using shelving pre- and post-filters," in *Acoustics, Speech and Signal Processing. ICASSP 2008. IEEE International Conference on*, Apr. 2008, pp. 209–212.

[13] J. Makhoul, "Stable and efficient lattice methods for linear prediction," Oct. 1977, vol. 25, pp. 423–428.

[14] R. Reininger and J. Gibson, "Backward adaptive lattice and transversal predictors for ADPCM," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84.*, Mar. 1984, vol. 9, pp. 429–432.

[15] L. Griffiths, "A continuously-adaptive filter implemented as a lattice structure," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '77.*, May 1977, vol. 2, pp. 683–686.

[16] Martin Holters, Christian R. Helmrich, and Udo Zölzer, "Delay-free audio coding based on ADPCM and error feedback," in *Proc. of the 11th Int. Conference on Digital Audio Effects (DAFx-08)*, Sep. 2008.

[17] D. Cohn and J. Melsa, "The relationship between an adaptive quantizer and a variance estimator (corresp.)," *Information Theory, IEEE Transactions on*, vol. 21, no. 6, pp. 669 – 671, Nov. 1975.

[18] International Telecommunication Union, "ITU recommendation ITU-R BS.1387-1, method for objective measurements of perceived audio quality (PEAQ)," Nov. 2001.

[19] Belgium European Broadcasting Union, Bruxelles, "Sound quality assessment material: Recordings for subjective tests – CD and user's handbook for the EBU-SQAM compact disc," Apr. 1988, Technical Centre of the EBU.

[20] Audio examples, "Error resilience enhancement of robust ADPCM audio coding," Oct. 2013, Available at http://ant.hsu-hh.de/conf/eradpcm/.