ASYMPTOTIC CLOSED-LOOP DESIGN OF ERROR RESILIENT PREDICTIVE COMPRESSION SYSTEMS

Sina Zamani, Tejaswi Nanjundaswamy, Kenneth Rose

Department of Electrical and Computer Engineering, University of California Santa Barbara, CA 93106
E-mail: {sinazmn, tejaswi, rose}@ece.ucsb.edu

ABSTRACT

Prediction is used in virtually all compression systems. When such a compressed signal is transmitted over unreliable networks, packet losses can lead to significant error propagation through the prediction loop. Despite this, the conventional design technique completely ignores the effect of packet losses, and estimates the prediction parameters to minimize the mean squared prediction error, and optimizes the quantizer to minimize the reconstruction error at the encoder. While some design techniques have been proposed to accurately estimate and minimize the end-to-end distortion at the decoder that accounts for packet losses, they operate in a closed-loop, which introduces a mismatch between statistics used for design and statistics used in operation, causing a negative impact on convergence and stability of the design procedure. Instead, we propose in this paper an effective technique for predictive compression system design that accounts for the instability caused by error propagation due to packet losses, and enjoys stable statistics during design by employing open-loop iterations that on convergence mimic closed-loop operation. Simulation results for a compression system with a first order linear predictor demonstrate the utility of the proposed approach, which offers significant performance improvements over existing design techniques.

Index Terms— Predictor design, Asymptotic Closed-Loop design, Error Resilience, Linear Predictor

1. INTRODUCTION

Linear prediction is widely used in speech coding, speech synthesis, speech recognition, audio coding, and video coding. In compression systems, the prediction module plays an important role in exploiting temporal and spatial redundancies. However, when such a compressed data is transmitted over unreliable networks, errors introduced due to the inevitable packet losses, propagate through the prediction loop, causing substantial, and sometimes catastrophic, deterioration of the received signal. Despite this, conventional compression system design completely ignores the effect of channel loss, and chooses the prediction parameters to minimize the mean squared prediction error, and optimizes the quantizer to minimize the reconstruction error at the encoder. This problem can be alleviated by optimizing the system for the overall end-to-end distortion (EED) observed at the decoder, which accounts for the effect of packet loss. An optimal recursive technique to estimate EED at the encoder was proposed in [1], and utilizing this distortion to optimally select the parameters for motion compensated prediction in video coding was proposed in [2]. However, designing optimal predictors and quantizers, while accounting for EED is a challenging task, as we need to work with a stable training set that accurately represents the true signal statistics. The open-loop (OL) and closed-loop (CL) approaches were proposed in [3] for predictive vector quantization and have been widely used since then. The OL approach uses the original data as the prediction reference during design, but since the decoder does not have access to the original data, the parameters designed are not suitable for the statistics seen at the decoder. The CL approach attempts to alleviates the problem of this mismatch by designing the parameters using reconstructed data obtained in a closed-loop system as the prediction reference. However, using these parameters in a closed-loop system generates new prediction and reconstruction, which implies they differ from the data the parameters were designed for. This mismatch in statistics between design and operation grows over time as the data is fed through the prediction loop in the coder, leading to instability in both estimation of prediction parameters, and design of quantizers, especially at lower bit rates. Note that this error propagation encountered during the design phase due to statistics mismatch, differs from the error propagation due to packet losses.

In this paper we propose to address the challenging problem of tackling these two types of error propagation, by designing the system iteratively, wherein an estimated EED is minimized at each iteration to account for packet losses, and the prediction reference from the previous iteration is employed in an open-loop way to ensure statistics used for design and operation are matched. Once the parameters being designed converge, the prediction reference in the current iteration will match the reference from the previous iteration, thus mimicking closed-loop operation. Hence, we call this the asymptotic closed-loop (ACL) approach, which is similar to the approach in [4, 5], wherein system design without accounting for packet losses is proposed. We specifically describe a framework for rate versus EED optimization of a compression system employing a first order linear predictor. We also propose a new encoder architecture, in which the prediction at encoder is based on the expected decoder reconstructions. Experimental results substantiate the utility of the proposed approach with significant performance improvements over existing design techniques.

2. PROBLEM SETUP

Fig. 1 illustrates a predictive compression system, wherein input signal samples, $x_n,\ 0 \leq n < N$, are coded by the encoder to generate a bitstream, which is transmitted through a channel to the decoder, where it is decoded to generate the reconstructed samples. The encoder uses its previous reconstructed samples, $\hat{x}_{e,n}$ to generate predicted samples, $\tilde{x}_{e,n}$ and the prediction error, $e_n = x_n - \tilde{x}_{e,n}$. This is quantized to generate $\hat{e}_n$, and sent over the channel. When the decoder receives $\hat{e}_n$, it adds it to its predicted sample, $\tilde{x}_{d,n}$ to generate its reconstructed samples, $\hat{x}_{d,n}$. Note that the reconstructed samples

This work was supported in part by a gift from Mozilla.

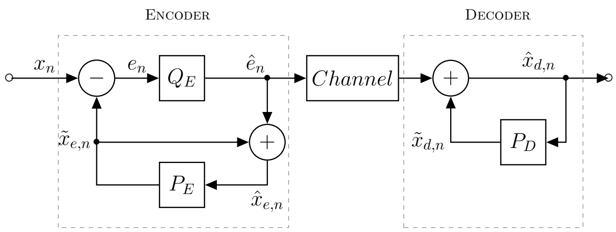


Fig. 1: A predictive compression system

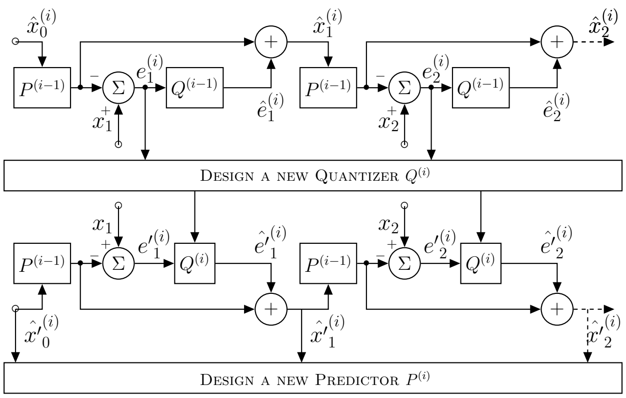


Fig. 2: Closed-loop training approach

at the encoder, $\hat{x}_{e,n}$, and the decoder, $\hat{x}_{d,n}$, will differ when the channel is unreliable and packets carrying $\hat{e}_n$ are lost. This uncertainty results in $\hat{x}_{d,n}$ being a random variable to the encoder. The problem at hand is to design optimal quantizers and predictors (P_E, Q_E and P_D) to minimize the expected EED at the decoder to account for packet losses. For the mean squared error distortion metric, expected EED at the decoder is,

$$\begin{aligned} E\{D\} &= \sum_{n=0}^{N-1} E\{(x_n - \hat{x}_{d,n})^2\} \\ &= \sum_{n=0}^{N-1} x_n^2 - 2x_n E\{\hat{x}_{d,n}\} + E\{(\hat{x}_{d,n})^2\}. \end{aligned} \tag{1}$$

Clearly, to estimate this distortion, first and second moments of the decoder reconstructions should be accurately estimated at the encoder.

3. BACKGROUND

3.1. End to End distortion estimation and prediction

A recursive technique to optimally estimate the expected EED at the encoder in the presence of packet losses via the first and second moments of the decoder reconstructions was proposed in [1] for video coders. The recursive algorithm optimally estimates the decoder reconstructions' first moment, $E\{\hat{x}^j_{d,n}\}$, and the second moment, $E\{(\hat{x}^j_{d,n})^2\}$, for every pixel j in frame n. These moments are then used to estimate EED at the encoder using (1) to optimally switch between inter-frame prediction and intra-frame prediction, to control the error propagation through frames. In [2], a new prediction scheme is employed in conjunction with optimal EED estimation.

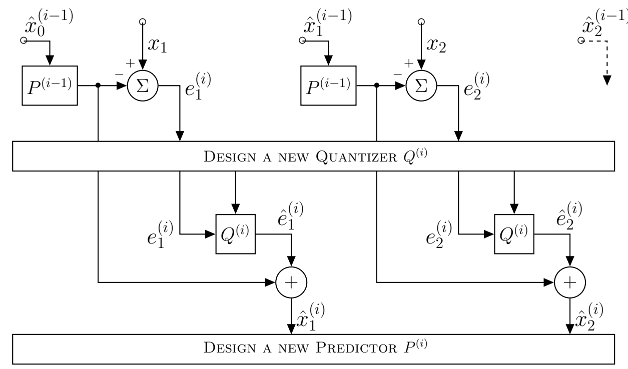


Fig. 3: Asymptotic closed-loop training approach

Conventional motion compensated prediction employs the encoder reconstructions for prediction, i.e., $\tilde{x}^j_{e,n} = \hat{x}^{j+v}_{e,n-1}$, where v is the optimal motion vector that minimizes the prediction error. Instead in [2], the prediction is based on expected decoder reconstructions, i.e., $\tilde{x}^j_{e,n} = E\{\hat{x}^{j+v}_{d,n-1}\}$, where v is the optimal motion vector that minimizes the EED of (1). This setup plays an important role in limiting error propagation during decoder operation by appropriately selecting motion vectors to predict from reference blocks that are less likely to be corrupted by error propagation.

3.2. Closed-Loop versus Asymptotic Closed-Loop Design

In closed-loop iterative design [6], the coder operates in closed-loop at each iteration to generate prediction errors and reconstructed samples that are used to design the updated quantizer and the updated predictor, respectively. At iteration $i-1$, given a quantizer, $Q^{(i-1)}$, and a predictor, $P^{(i-1)}$, a training set of prediction errors, $T^{(i)} : \{e_n^{(i)}\}_{n=1}^N$, for iteration i is generated as,

$$e_n^{(i)} = x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i)}), \tag{2}$$

where,

$$\hat{x}_n^{(i)} = P^{(i-1)}(\hat{x}_{n-1}^{(i)}) + Q^{(i-1)}(x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i)})). \tag{3}$$

These two equations are calculated sequentially for all values n. Then given $T^{(i)}$, we design a new quantizer, $Q^{(i)}$. Using $Q^{(i)}$, a new set of reconstructed samples, $\hat{x'}_n^{(i)}$, is generated as per (3) and based on this, we design a new predictor, $P^{(i)}$. These steps are repeated until convergence. Fig. 2 illustrates this closed-loop iterative design. The major issue with this approach is that when the updated parameters are employed in closed-loop at an iteration, new prediction errors are generated, which differ from the errors the quantizer was designed for, and this implies different reconstructions are generated, which differ from the reference reconstructions the predictor was designed for. This mismatch in statistics between design and operation builds up over time as the data is fed through the prediction loop in the coder, leading to instability in the iterative design of both the predictor and the quantizer, especially at lower bit rates. The ACL design technique proposed in [4, 5], tackles this statistics mismatch issue by designing the predictor and the quantizer in an open-loop fashion, while ultimately optimizing the system for closed-loop operation. Specifically, the prediction is based on reconstructions of previous iteration, i.e.,the prediction errors are generated as,

$$e_n^{(i)} = x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i-1)}). \tag{4}$$

Given the new prediction errors, we design a new quantizer, $Q^{(i)}$. This $Q^{(i)}$ is now employed to generate the reconstructed samples of next iteration as,

$$\hat{x}_n^{(i)} = P^{(i-1)}(\hat{x}_{n-1}^{(i-1)}) + Q^{(i)}(x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i-1)})), \quad (5)$$

again using the reconstructions of previous iteration for prediction. Given these new reconstructions, we design a new predictor, $P^{(i)}$. Note that the equations (4) and (5) are executed independently for each sample of the sequence in an open-loop way. The main steps of this technique are depicted in Fig. 3. The open-loop format ensures the predictor and quantizer employ exactly the same reconstructed data and prediction error used for their design, eliminating the statistical mismatch issue seen in closed-loop design. On convergence, the predictor and the quantizer do not change, which implies, $\hat{x}_{n-1}^{(i)} = \hat{x}_{n-1}^{(i-1)}$, i.e., predicting from previous iteration reconstructions is the same as predicting from the current iteration reconstructions, which is effectively closed-loop operation.

4. PROPOSED APPROACH

We propose a framework to design a first order predictor and a quantizer to minimize the EED in (1). We first develop the EED estimation algorithm, then we propose an encoder architecture in which predictions are based on the expected reconstructions at the decoder and finally we propose the ACL design approach that accounts for packet loss to improve coding efficiency and design stability.

4.1. Expected Decoder Distortion and Reconstructions

We assume for simplicity of presentation that each packet contains one sample (or alternatively that interleaving is used). The packet (or sample) loss rate is denoted as p. The prediction model employed at the decoder is a simple first order linear predictor,

$$\tilde{x}_{d,n} = \alpha \hat{x}_{d,n-1}, \quad (6)$$

where α is the prediction coefficient that needs to be estimated. The quantized prediction error, $\hat{e}_n$, transmitted over the channel, may or may not be received by the decoder. If the current packet is received (with probability $1-p$), the decoder uses it to generate the reconstructed sample as,

$$\hat{x}_{d,n} = \tilde{x}_{d,n} + \hat{e}_n. \quad (7)$$

When the packet is lost, a simple concealment of setting residue to zero is employed, which gives the reconstructed sample as,

$$\hat{x}_{d,n} = \tilde{x}_{d,n}. \quad (8)$$

Thus the first and second moment of the decoder reconstructed samples, required to estimate EED given in (1), are calculated recursively at the encoder as,

$$\begin{aligned} E\{\hat{x}_{d,n}\} &= (1-p)E\{\hat{e}_n + \alpha\hat{x}_{d,n-1}\} + pE\{\alpha\hat{x}_{d,n-1}\} \\ &= (1-p)\hat{e}_n + \alpha E\{\hat{x}_{d,n-1}\} \end{aligned} \quad (9)$$

$$\begin{aligned} E\{(\hat{x}_{d,n})^2\} &= (1-p)E\{(\hat{e}_n + \alpha\hat{x}_{d,n-1})^2\} + pE\{(\alpha\hat{x}_{d,n-1})^2\} \\ &= (1-p)(\hat{e}_n^2 + 2\alpha\hat{e}_n E\{\hat{x}_{d,n-1}\}) + \alpha^2 E\{(\hat{x}_{d,n-1})^2\} \end{aligned} \quad (10)$$

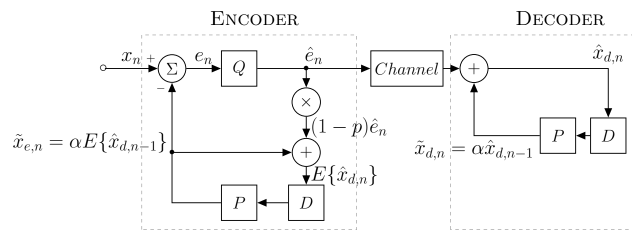


Fig. 4: Architecture of the proposed coder

4.2. Prediction Based on the Expected Decoder Reconstructions

Packet losses cause the reconstructions at the encoder and the decoder to differ. Thus to close the gap between prediction at the encoder and the decoder, we employ the expected decoder reconstructions for prediction at the encoder, i.e.,

$$\tilde{x}_{e,n} = \alpha E\{\hat{x}_{d,n-1}\}. \quad (11)$$

The prediction error, $e_n = x_n - \tilde{x}_{e,n}$, is then quantized to generate, $\hat{e}_n$. The overall proposed architecture is shown in Fig. 4.

We design the prediction coefficient α to minimize the EED, by solving for α in the equation given by setting the partial derivative of EED with respect to α to 0. The EED in (1) is dependent on α through equations (9) and (10). The equation to be solved is,

$$\begin{aligned} \frac{\partial E\{D\}}{\partial \alpha} &= \sum_{n=0}^{N-1} -2x_n E\{\hat{x}_{d,n-1}\} + \\ &\quad \sum_{n=0}^{N-1} 2(1-p)\hat{e}_n E\{\hat{x}_{d,n-1}\} + 2\alpha E\{(\hat{x}_{d,n-1})^2\} \\ &= 0, \end{aligned} \quad (12)$$

which gives us the solution as,

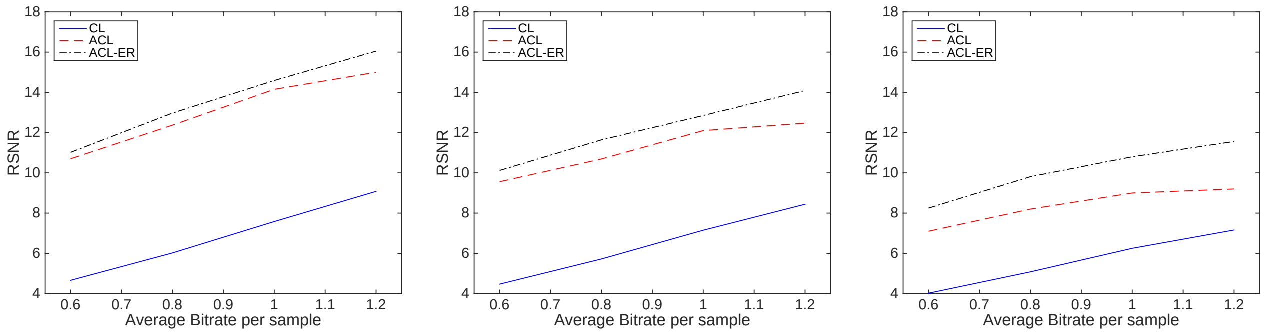
$$\alpha = \frac{\sum_{n=0}^{N-1} E\{\hat{x}_{d,n-1}\}(x_n - (1-p)\hat{e}_n)}{\sum_{n=0}^{N-1} E\{(\hat{x}_{d,n-1})^2\}}. \quad (13)$$

Note that although $\hat{e}_n$ is dependent on α, we assume that the modifications in α across our design iterations are small enough to not change the quantization intervals.

4.3. Asymptotic Closed-Loop Design

We employ the ACL approach for a stable system design by eliminating the statistical mismatch issue of closed-loop design. This is achieved by operating in an open-loop way, wherein we employ previous iteration's first and second moments of the decoder reconstructions, to estimate current iteration's moments and prediction. Given a set of decoder reconstructions' first moments, $E\{\hat{x}_d\}^{(i-1)}$, and second moments, $E\{(\hat{x}_d)^2\}^{(i-1)}$, of iteration $i-1$, the predictor and quantizer are iteratively designed in an inner loop. In a subiteration s of the inner loop, given a set of quantized prediction errors, $\hat{e}_n^{(i,s-1)}$, the optimal prediction coefficient is estimated as,

$$\alpha^{(i,s)} = \frac{\sum_{n=0}^{N-1} E\{\hat{x}_{d,n-1}\}^{(i-1)}(x_n - (1-p)\hat{e}_n^{(i,s-1)})}{\sum_{n=0}^{N-1} E\{(\hat{x}_{d,n-1})^2\}^{(i-1)}}. \quad (14)$$



(a) (b) (c)

Fig. 5: Average decoder distortion (in dB) of CL, ACL and the proposed ACL-ER design approaches, for a first order predictive coder with entropy constrained scalar quantizer, at various bit rates for packet loss rates of (a)5% (b)10% (c)20%.

The new prediction errors are now generated in an open-loop fashion as,

$$e_n^{(i,s)} = x_n - \alpha^{(i,s)} E\{\hat{x}_{d,n-1}\}^{(i-1)}. \quad (15)$$

An optimal quantizer, $Q^{(i,s)}$, is now designed for this set of new prediction errors, which is used to generate a new set of quantized prediction errors, $\hat{e}_n^{(i,s)} = Q^{(i,s)}(e_n^{(i,s)})$. These subiterations are repeated until convergence to obtain current subiterations' final quantizer, $Q^{(i)}$, final prediction coefficient, $\alpha^{(i)}$, and final set of quantized prediction errors, $e_n^{(i)}$. The first and second moments of the decoder reconstructions are now updated in the outer loop in an open-loop way as,

$$E\{\hat{x}_{d,n}\}^{(i)} = (1-p)\hat{e}_n^{(i)} + \alpha^{(i)} E\{\hat{x}_{d,n-1}\}^{(i-1)} \quad (16)$$

$$E\{(\hat{x}_{d,n})^2\}^{(i)} = (1-p)((\hat{e}_n^{(i)})^2 + 2\alpha^{(i)}\hat{e}_n^{(i)} E\{\hat{x}_{d,n-1}\}^{(i-1)}) + (\alpha^{(i)})^2 E\{(\hat{x}_{d,n-1})^2\}^{(i-1)}. \quad (17)$$

These moments are now used in the next iterations inner loop to update the predictor and quantizer. Iterations are repeated until convergence. Note that although the entire design is in open-loop, on convergence it emulates closed-loop operation. This is achieved as on convergence the quantizer and predictor do not change, i.e., $Q^{(i)} = Q^{(i-1)}$ and $\alpha^{(i)} = \alpha^{(i-1)}$, which implies $E\{\hat{x}_{d,n}\}^{(i)} = E\{\hat{x}_{d,n}\}^{(i-1)}$, thus employing previous iteration's moments is the same as estimating current iteration's moments recursively and employing them for prediction in a closed-loop way.

5. EXPERIMENTAL RESULTS

To validate our proposed method, we evaluated it for a compression system with first order linear prediction and an entropy constrained scalar quantizer. The Generalized Lloyd Algorithm (GLA) was used to design the entropy constrained scalar quantizer. We used the 6 speech files available in the EBU SQAM database [7] as our dataset, as linear prediction is commonly employed in speech coding. However, note that the proposed approach is applicable to predictive compression of any signal with temporal correlations. The first half of the speech files were used as training set (resulting in more 2 million samples) and the second half as test data. The prediction coefficient was initialized to zero. We evaluated the following three different design techniques:

1. The closed-loop design procedure discussed in Section 3.2, which completely ignores the packet losses (referred as CL).
2. The ACL algorithm discussed in [5], which also ignores the packet losses and can be obtained by setting $p = 0$ in the update formulas of Section 4.3 (referred as ACL).
3. Our proposed method (referred as ACL-ER).

The system performance is evaluated by plotting the decoder reconstruction error's signal to noise ratio (RSNR) averaged over ten different loss patterns versus average bitrate, as shown in Fig. 5 for different packet loss rates. Clearly, the proposed approach consistently outperforms both ACL and CL under all testing scenarios, with gains of up to 7 dB over CL and gains of up to 2.5 dB over ACL. The proposed approach provides higher performance improvements as the packet loss rate increases, since accounting for error propagation due to packet losses becomes critical in these cases. Compared to ACL, our method provides larger gains at higher bitrates, as ACL relies more on the high quality previous reconstructions for prediction, unaware of the fact that these reconstructions at the decoder will be corrupted as a result of error propagation due to packet losses. These results substantiate the significant utility of the proposed approach.

6. CONCLUSION

In this paper we proposed an effective and robust design technique for a predictive compression system. We account for the presence of an unreliable channel by designing the system to optimize the EED at the decoder. We then eliminate the statistical mismatch issue suffered by conventional closed-loop approaches, by employing a stable iterative design approach that operates in an open-loop way and on convergence mimics closed-loop operation. By carefully designing the system parameters, error propagation at the decoder is effectively contained. Significant performance improvements seen in experimental evaluation results demonstrate the utility of the proposed approach. Future research directions include, extending the proposed design technique to higher order predictors, and employing a powerful optimization technique for design of the entropy constrained scalar quantizer to account for packet losses.

7. REFERENCES

[1] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience,"

IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, pp. 966–976, 2000.

[2] H. Yang and K. Rose, "Source-channel prediction in error resilient video coding," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2003, vol. 2, pp. II–233.

[3] V. Cuperman and A. Gersho, "Vector predictive coding of speech at 16 kbits/s," *IEEE Transactions on Communications*, vol. 33, no. 7, pp. 685–696, 1985.

[4] H. Khalil, K. Rose, and S. L. Regunathan, "The asymptotic closed-loop approach to predictive vector quantizer design with application in video coding," *IEEE Transactions on Image Processing*, vol. 10, no. 1, pp. 15–23, 2001.

[5] H. Khalil and K. Rose, "Robust predictive vector quantizer design," in *Data Compression Conference (DCC)*, 2001, pp. 33–42.

[6] P.-C Chang and R. M. Gray, "Gradient algorithms for designing predictive vector quantizers," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 4, pp. 679–690, 1986.

[7] G Waters, "Sound quality assessment materialrecordings for subjective tests: Users handbook for the ebu–sqam compact disk," *European Broadcasting Union (EBU), Tech. Rep*, 1988.