DETERMINISTIC ANNEALING BASED DESIGN OF ERROR RESILIENT PREDICTIVE COMPRESSION SYSTEMS

Bharath Vishwanath, Tejaswi Nanjundaswamy, Sina Zamani and Kenneth Rose

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

ABSTRACT

This paper considers near optimal design of predictive compression system that accounts for packet loss over unreliable networks. Major challenges to address include, propagation of errors due to packet loss through the prediction loop, mismatch between statistics used for design and during operation, and above all a cost function that is fraught with poor local minima. Accurately estimating and minimizing the end-to-end distortion (EED), in combination with asymptotic closed-loop (ACL) design that employs open-loop iterations, but mimics closed-loop operation on convergence, was proposed to address the first two challenges. However the severe non-convexity of the cost function, especially due to the piece-wise linear nature of the quantizer function, makes this a particularly challenging optimization problem. We propose to tackle this via a new design approach in the deterministic annealing framework to avoid poor local minima, coupled with the ACL approach to minimize EED estimate. This effectively addresses all the major challenges, and leads to a near optimal design of error-resilient predictive compression system. Substantial performance improvement obtained in experimental evaluations demonstrates the efficacy of the proposed approach.

Index Terms— Predictor design, annealing, error resilience, asymptotic closed-loop

1. INTRODUCTION

Prediction is widely used in speech, audio and video coding to exploit spatial and temporal correlations [1]-[3]. These coders predict the current sample from previously reconstructed samples, then quantize and transmit the prediction error. However, when such compressed data is transmitted over unreliable networks, packet loss in networks leads to error in reconstructed signal. This error further propagates in the prediction loop and can lead to significant deterioration in the reconstructed signal. In spite of this, traditional approaches do not account for the packet loss, and design predictors and quantizers to simply minimize the distortion at the encoder. Various error control and concealment approaches were proposed in [4] and redundancy coding to improve robustness was proposed in [5], [6]. However, in order to build robust systems that account for channel loss, we need to optimize the end-to-end distortion (EED) observed at the decoder. Optimal estimation of EED was proposed in [7] and a predictor designed to minimize EED was proposed in [8].

Still, joint design of the predictor and quantizer of a predictive compression system is further challenging as the feedback loop creates a complex relation between the two. Two simple design approaches of, open-loop and closed-loop were introduced in [9]
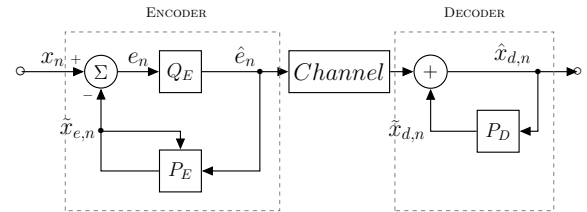
Fig. 1. Predictive compression system

and [10], respectively, which minimize the distortion at the encoder. These approaches suffer from stability issues, especially at low bit-rates, due to mismatch between statistics used for design and statistics observed during operation. To address this, the asymptotic closed-loop (ACL) design was proposed in [11]. ACL iteratively updates the predictor and quantizer in open-loop to ensure stability of statistics used during design and operation, but on convergence mimics closed-loop operation, thus enjoying benefits of both open-loop and closed-loop design. To account for packet loss, ACL based design of predictive compression system to minimize EED estimate was proposed in [12]. This approach addressed the challenges of avoiding mismatch of statistics during design and operation, and accounting for error propagation due to packet loss. However, the severe non-convexity of the cost function, especially given the piece-wise linearity of the quantizer, makes it extremely challenging to find an optimal solution for the joint design of predictor and quantizer. That is, the current design approaches cannot sense the global minimum and can easily get trapped in poor local minima.

In this paper, we present a new framework which tackles all the challenges in design of error-resilient predictive compression system to achieve a near optimal solution. Inspired by the principles of statistical physics and information theory, deterministic annealing (DA) was proposed as a powerful non-convex optimization tool in [13]. DA offers two advantages: its probabilistic formulation replaces piece-wise linear quantization with a differentiable cost function that can be jointly optimized for the predictor and quantizer parameters, and its annealing schedule avoids many poor local minima of the cost surface. While DA based ACL design was proposed in [14], the impact of packet loss was not accounted. Instead, we propose a new approach that embeds ACL design in DA framework to minimize EED at each iteration. Substantial gains in simulations validates the utility of the proposed approach.

2. PROBLEM FORMULATION

Fig. 1 shows a predictive compression system. Let x_n, $0 \le n \le N$ be the input samples. Encoder predicts a sample x_n based on pre-

viously reconstructed samples as $\tilde{x}_{e,n}$. The prediction error, $e_n = x_n - \tilde{x}_{e,n}$, is quantized to $\hat{e}_n$. These quantized values are communicated to the decoder over the channel. At the decoder, the quantized error is added to the predicted value $\tilde{x}_{d,n}$ to get the reconstructed sample $\hat{x}_{d,n}$. When there is a packet loss, the decoder does error concealment and thus the decoder reconstructed samples $\hat{x}_{d,n}$ differ from the encoder reconstructed samples $\hat{x}_{e,n}$. We need to minimize the distortion as experienced by the decoder. Since the channel is not deterministic, the decoder reconstructed sample is a random variable to the encoder. This means the encoder can only estimate the distortion as seen at the decoder. With squared error as the distortion metric, we can calculate the expected EED as,

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

It is evident that in order to estimate EED, we need to accurately estimate the first and second order moments of the decoder reconstruction. The overall problem is to design (P_E, P_D, Q_E) to minimize $E\{D\}$.

3. BACKGROUND

3.1. EED Estimation

A recursive method to optimally estimate the expected decoder distortion per pixel in video coding was proposed in [7]. The decoder distortion was estimated by recursively computing the first and second order moments of the decoder reconstruction. The estimated EED was used to optimally switch between inter and intra modes for coding a macroblock, limiting the error propagation due to packet loss. To minimize the expected EED, a new prediction framework was proposed in [8], which used the expected decoder reconstruction for the prediction as against to the traditional approach of using the reconstructed samples at the encoder.

3.2. Open-loop, closed-loop and asymptotic closed-loop design

In the open-loop approach [9], the predictor and quantizer are designed using the original source samples. Predictor, P, is designed based on correlation of original samples and the quantizer training error is also generated using the original samples, i.e., $e_n = x_n - P(x_{n-1})$. Although this approach provides a stable training set of prediction errors, it suffers from the serious shortcoming that the decoder doesn't have access to original samples. In the closed-loop approach [10], the predictor and quantizer are designed iteratively. In iteration i, given previous iteration predictor, $P^{(i-1)}$, and quantizer, $Q^{(i-1)}$, a training set of errors,

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

is generated, 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 errors are used to design a new quantizer, $Q^{(i)}$. Then given $Q^{(i)}$, and $P^{(i-1)}$, new reconstructed samples,

$$
\hat{x}'_n^{(i)} = P^{(i-1)}(\hat{x}'_{n-1}^{(i)}) + Q^{(i)}(x_n - P^{(i-1)}(\hat{x}'_{n-1}^{(i)})),
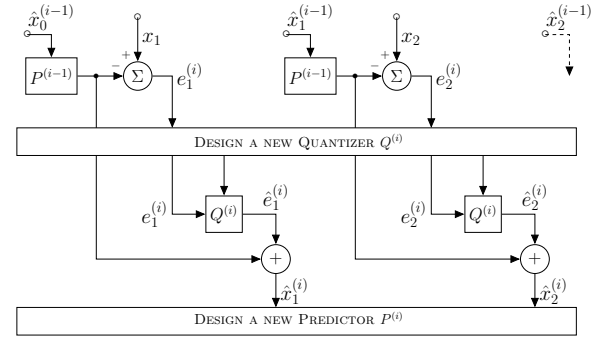\tag{4}
$$

Fig. 2. Asymptotic closed loop design

are generated and used to design a new predictor, $P^{(i)}$. Notice in (4), $Q^{(i)}$ is used with $(x_n - P^{(i-1)}(\hat{x}'_{n-1}^{(i)}))$, while it was designed for $(x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i)}))$. Similarly, $P^{(i)}$ is designed for $\hat{x}'_n^{(i)}$, but used in the next iteration with $\hat{x}_n^{(i+1)}$. This mismatch in statistics between design and operation, builds over time through the prediction loop, leading to instability of the iterative closed-loop design, especially for low bitrates. This instability problem was addressed by the ACL approach [11], which is built based on best of both the worlds. At each iteration, prediction is based on the reconstructed samples of the previous iteration and the training set is generated in an open loop fashion, i.e, $e_n^{(i)} = x_n - P^{(i-1)}(\hat{x}_{n-1}^{(i-1)})$. A new quantizer, $Q^{(i)}$, is trained using these errors, and then used to quantize the same errors, and generate reconstructed samples,

$$
\hat{x}_n^{(i)} = P^{(i-1)}(\hat{x}_{n-1}^{(i-1)}) + Q^{(i)}(e_n^{(i)}).
\tag{5}
$$

These samples are now used to train a new predictor, $P^{(i)}$. The ACL approach is illustrated in Fig. 2, and one can clearly observe that the quantizer and predictor are used for exactly the same statistics they were designed for. Upon convergence, the reconstructed samples remain the same over iterations. Thus, predicting from $\hat{x}_{n-1}^{(i-1)}$ is same as predicting from $\hat{x}_{n-1}^{(i)}$, which is essentially closed-loop operation.

3.3. ACL-EED

Design of a predictive compression system via ACL approach to minimize EED estimate was proposed in [12]. For simplicity of presentation, a first order prediction model was assumed. Thus the decoder prediction is, $\tilde{x}_{d,n} = \alpha\hat{x}_{d,n-1}$. Also each packet was assumed to carry one quantized error sample. If the packet is received, the quantized error is added to the decoder prediction, i.e, $\hat{x}_{d,n} = \tilde{x}_{d,n} + \hat{e}_n$. In the event of packet loss, the residue is considered zero. With this framework and a packet loss rate of p, the first and second order moments were recursively estimated to be,

$$
\begin{aligned}
E\{\hat{x}_{d,n}\} &= (1-p)E\{(\alpha\hat{x}_{d,n-1} + \hat{e}_n)\} + pE\{\alpha\hat{x}_{d,n-1}\} \\
&= \alpha E\{\hat{x}_{d,n-1}\} + (1-p)\hat{e}_n
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
E\{(\hat{x}_{d,n})^2\} &= (1-p)E\{(\alpha\hat{x}_{d,n-1} + \hat{e}_n)^2\} + pE\{(\alpha\hat{x}_{d,n-1})^2\} \\
&= \alpha^2 E\{(\hat{x}_{d,n-1})^2\} + (1-p)(\hat{e}_n^2 + 2\alpha E\{\hat{x}_{d,n-1}\}\hat{e}_n)
\end{aligned}
\tag{7}
$$

Similar to [8], the prediction at the encoder was based on the expected decoder reconstruction i.e, $\tilde{x}_{e,n} = \alpha E(\hat{x}_{d,n-1})$. The optimal

predictor which minimizes the expected EED was derived to be,

$$\alpha = \frac{\sum_{n=0}^{N} E\{\hat{x}_{d,n-1}\}(x_n - (1-p)\hat{e}_n)}{\sum_{n=0}^{N} E\{(\hat{x}_{d,n-1})^2\}}. \qquad (8)$$

Asymptotic closed loop design in two nested loops was employed to obtain optimal predictor and quantizer. In an iteration of the outer loop, decoder moments were fixed, and the predictor and quantizer were optimized in the inner loop in the ACL way. The decoder moments were then updated, again in the ACL way, using previous moments from the previous iteration. Upon convergence, the operation mimicked closed-loop.

4. PROPOSED APPROACH

The piece-wise linear nature of the quantizer makes it difficult to optimize a joint cost function of the predictor and quantizer, since the derivative vanishes almost everywhere. Hence the iterative approach was employed before. However, this only ensures optimality of the quantizer or the predictor separately for the local training set in an iteration. This approach naturally does not guarantee global optimality, and can easily get trapped in poor local minima. We thus propose a novel approach of coupling the ACL based minimization of expected EED with the DA framework to overcome the suboptimality of the prior technique. The proposed approach builds on DA introduced in [13], which is based on the intuition gained from annealing process in physical chemistry, where certain systems are driven to their low energy states by gradually cooling the system. Analogously, we introduce controlled randomness in the optimization procedure, but deterministically minimize the expected EED, thereby avoiding many poor local minima. The amount of randomness is measured by Shannon's entropy and is essentially characterized by the temperature of the system. The encoding rule is no longer piecewise linear, and is differentiable everywhere paving our way to jointly optimize predictor and quantizer. It can be easily shown that quantizer distortion given a training set is independent of EED. With $\{y_j\}$ being the codebook and $P_{j|n}$ being the association probability of e_n to codevector y_j, similar to the formulation in [15], the total squared error distortion which is minimized in the entropy constrained quantizer design in an iteration i of ACL will now take probabilistic version as,

$$J = \sum_n \sum_j P_n^i P_{j|n}^i \left\{ (e_n^i - y_j^i)^2 - \lambda \log(P_j^i) \right\}, \qquad (9)$$

where, P_n^i is the probability of training vector. The randomness is measured by Shannon's entropy as,

$$H = -\sum_n \sum_j P_{nj}^i \log(P_{nj}^i), \qquad (10)$$

where, $P_{nj}^i = P_n^i P_{j|n}^i$, is the joint probability distibution of the codevector and the training sample. With the degree of randomness measured by Shannon's entropy and the distortion as defined above, the problem at hand can be viewed as minimization of the Lagrangian cost function,

$$F = J - TH. \qquad (11)$$

The degree of randomness is controlled by T, which characterizes the temperature. Assuming uniform distribution over the training set, it can been shown that minimizing the lagrangian cost with respect
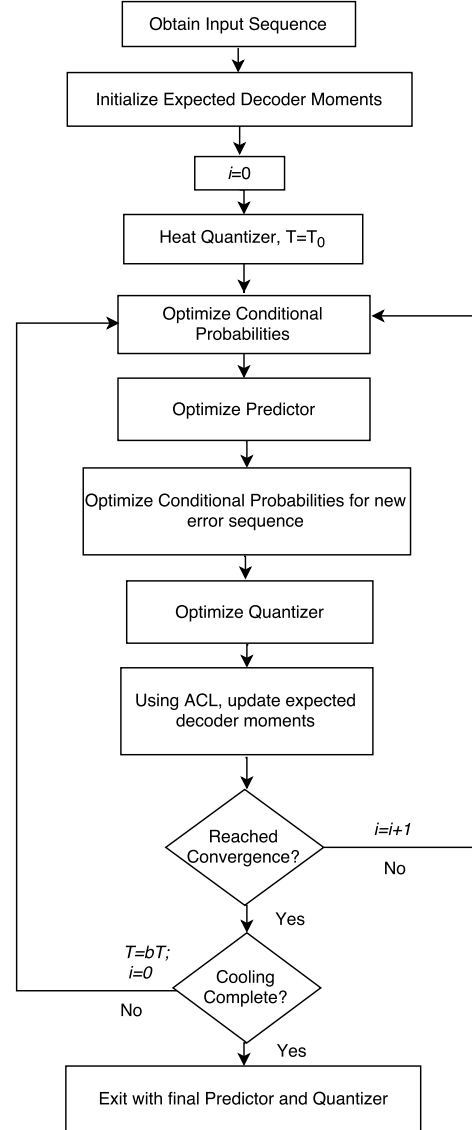

Fig. 3. Flow chart of the proposed algorithm

to the association probabilities leads to Gibb's distribution for the association probabilities:

$$P_{j|n}^i = \frac{e^{-\{(e_n^i - y_j^i)^2 - \lambda \log(P_j^i)\}}}{\sum_k e^{-\{(e_n^i - y_k^i)^2 - \lambda \log(P_k^i)\}}}. \qquad (12)$$

It is to be observed that at high temperatures, we in fact maximize entropy of the system and thus the association probabilities are uniform. As we lower the temperature, we trade entropy for distortion and the system becomes more deterministic. In the probabilistic framework, the recursive calculation of the decoder moments now becomes,

$$E\{\hat{x}_{d,n}^i\} = \alpha^i E\{\hat{x}_{d,n-1}^{i-1}\} + (1-p)\sum_j P_{j|n}^i y_j^i$$

$$E\{(\hat{x}_{d,n}^i)^2\} = (\alpha^i)^2 E\{(\hat{x}_{d,n-1}^{i-1})^2\} + (1-p)\sum_j P_{j|n}^i \{(y_j^i)^2 + 2\alpha E\{\hat{x}_{d,n-1}^{i-1}\}y_j^i\}. \qquad (13)$$

The optimal predictor now satisfies,

$$
\begin{aligned}
\frac{\partial E\{D\}}{\partial \alpha^i} &= -2\sum_n x_n E\{\hat{x}_{d,n-1}^{i-1}\} + \sum_n 2\alpha^i E\{(\hat{x}_{d,n-1}^{i-1})^2\} + \\
&\quad \sum_n \sum_j 2(1-p)P_{j|n}^i y_j^i E\{\hat{x}_{d,n-1}^{i-1}\} \\
&= 0. \qquad\qquad\qquad\qquad\qquad\qquad (14)
\end{aligned}
$$

Hence,

$$
\alpha^i = \frac{\sum_{n=0}^N E\{\hat{x}_{d,n-1}^{i-1}\}(x_n - (1-p)\sum_j P_{j|n}^i y_j^i)}{\sum_{n=0}^N E\{(\hat{x}_{d,n-1}^{i-1})^2\}}. \qquad (15)
$$

The optimal quantizer at iteration i satisfies,

$$
\frac{\partial}{\partial y_j} \sum_n \sum_j P_n^i P_{j|n}^i((e_n^i - y_j^i)^2 - \lambda \log(P_j^i)) = 0. \qquad (16)
$$

Hence,

$$
y_j^i = \frac{\sum_n P_{j|n}^i e_n^i}{\sum_n P_{j|n}^i}. \qquad (17)
$$

The overall proposed algorithm is illustrated as a flow chart in Fig. 3.

5. SIMULATION RESULTS

In our simulations, we compared the standard closed loop technique (CL) and the ACL-EED method proposed in [12] with the proposed approach (referred to as DA-ACL-EED). While the proposed design approach is applicable to any predictive compression system, speech files were considered for testing since linear prediction is commonly employed in speech coding. Among the set of 6 speech files from EBU SQAM database [16], 3 speech files were used to train the system and the remaining three were used for testing. The reconstructed SNR depends on the number of quantization levels for both ACL-EED and DA-ACL-EED. For fairness of comparison and to simplify simulations, the number of quantization levels was chosen to be four. The expected decoder moments and the prediction coefficient were initialized to zero. Simulations were carried out for different packet loss rates and the reconstructed SNR at the decoder was observed. In order to simulate the behavior of the channel, we randomly generated 20 packet loss patterns for each packet loss rate and averaged the reconstructed SNR. Fig. 4 shows the average reconstructed SNR for different packet loss rates and different entropy constraints. It can be seen that the current approach consistently performs better than CL and ACL-EED approaches with gains as high as 9.4 dB compared to CL and 2.7 dB compared to ACL-EED, and outperforms them significantly at low bitrates, highlighting the shortcomings of the prior approaches.

6. CONCLUSIONS

This paper describes a novel near optimal design approach for error-resilient predictive compression systems. The design iterations are open loop in nature but eventually converge to closed loop operation. The system is trained in a probabilistic way using deterministic annealing, avoiding poor local minima and simultaneously optimizing both predictor and quantizer. Expected EED is minimized to give parameters that are robust to channel loss. Substantial gains for different packet loss rates and entropy constraints demonstrates the utility of the new approach.
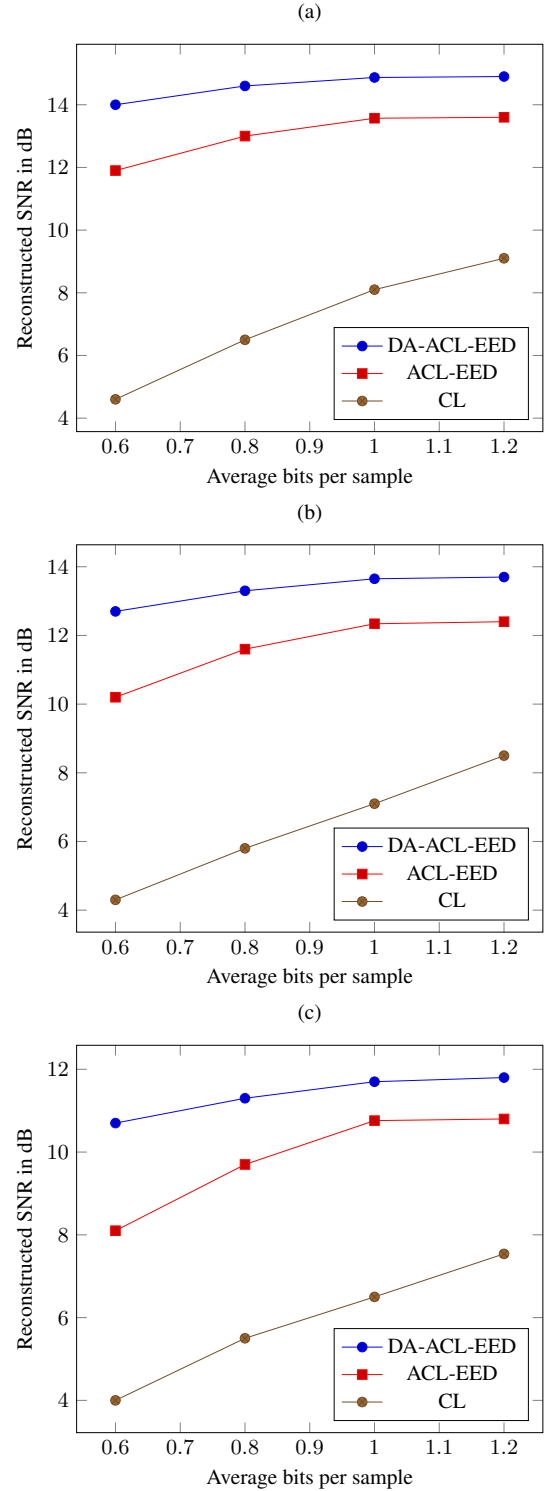

Fig. 4. Reconstructed SNR vs average bits per sample for packet loss rate (a) 5% (b) 10% and (c) 20%

7. REFERENCES

[1] A. S. Spanias, "Speech coding: a tutorial review," *Proceedings of the IEEE*, vol. 82, no. 10, pp.1541-1582, 1994.

[2] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 451-515, 2000.

[3] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.

[4] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: a review," *Proceedings of the IEEE*, vol. 86, p. 974-97, May 1998.

[5] G. Côté and F. I. Kossentini, "Optimal intra coding of blocks for robust video communication over the internet," *Signal Processing: Image Communication*, vol. 15, no. 1, pp. 25-34, 1999.

[6] S. Wenger, "Video redundancy coding in H.263+," in *Proc. AVSPN*, Aberdeen, U.K., Sept. 1997.

[7] 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.

[8] H. Yang and K. Rose, "Source-channel prediction in error resilient video coding," in *Proc. IEEE ICME*, 2003, vol. 2, pp. II-233.

[9] 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.

[10] 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.

[11] 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.

[12] S. Zamani, T. Nanjundaswamy, and K. Rose, "Asymptotic closed-loop design of error resilient predictive compression systems", in *Proc. IEEE ICASSP*, Mar. 2016.

[13] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2210-2239, 1998.

[14] H. Khalil and K. Rose, "Predictive vector quantizer design using deterministic annealing," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 244-254, Jan. 2003.

[15] K. Rose and D. Miller, "A deterministic annealing algorithm for entropy-constrained vector quantizer design," *27th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, November 1993.

[16] G. Waters, "Sound quality assessment material recordings for subjective tests: Users handbook for the EBU-SQAM compact disk," *European Broadcasting Union (EBU), Tech. Rep*, 1988.