

Efficient Bit-Rate Scalability for Weighted Squared Error Optimization in Audio Coding

Ashish Aggarwal, *Member, IEEE*, Shankar L. Regunathan, *Member, IEEE*, and Kenneth Rose, *Fellow, IEEE*

Abstract—We propose two quantization techniques for improving the bit-rate scalability of compression systems that optimize a weighted squared error (WSE) distortion metric. We show that quantization of the base-layer reconstruction error using entropy-coded scalar quantizers is suboptimal for the WSE metric. By considering the compandor representation of the quantizer, we demonstrate that asymptotic (high resolution) optimal scalability in the operational rate-distortion sense is achievable by quantizing the reconstruction error in the compandor's companded domain. We then fundamentally extend this work to the low-rate case by the use of enhancement-layer quantization which is conditional on the base-layer information. In the practically important case that the source is well modeled as a Laplacian process, we show that such conditional coding is implementable by only two distinct switchable quantizers. Conditional coding leads to substantial improvement over the companded scalable quantization scheme introduced in the first part, which itself significantly outperforms standard techniques. Simulation results are presented for synthetic memoryless Laplacian sources with μ -law companding, and for real-world audio signals in conjunction with MPEG AAC. Using the objective noise-mask ratio (NMR) metric, the proposed approaches were found to result in bit-rate savings of a factor of 2 to 3 when implemented within the scalable MPEG AAC. Moreover, the four-layer scalable coder consisting of 16-kb/s layers achieves performance close to that of the 64-kb/s nonscalable coder on the standard test database of 44.1-kHz audio.

Index Terms—AAC, audio coding, bit-rate scalability, embedded transmission, quantization.

I. INTRODUCTION

THE problem of efficient bit-rate scalability, or embedded coding, is an important one. Bit-rate scalability is a central requirement in many audio compression systems aimed at wireless and networking applications. A scalable bit stream allows the decoder to produce a coarse reconstruction if only a portion of the bit stream is received, and to improve the quality as more of the total bit stream is made available. Scalability is especially important in applications such as digital audio broadcasting and

multicast audio, which require simultaneous transmission over multiple channels of differing capacity. Further, a scalable bit stream provides robustness to packet loss for transmission over packet networks (e.g., over the Internet). A recent standard for scalable audio coding is MPEG-4 [1] which performs multi-layer coding [2] using Advanced Audio Coding (AAC) [3]–[6] modules.

Current state-of-the-art audio coders including AAC, AC3 [7], PAC [8], and ATRAC [9] rely on exploiting perceptual irrelevancy via *quantization noise shaping* [10]–[12]. Hence, perceptually motivated objective metrics for audio differ from the simple mean squared error (mse) distortion criterion. For example, variants of the noise-mask ratio (NMR) [13]–[16], a weighted squared error (WSE) measure, are widely employed as the distortion metric for encoder optimization. Quantization in perceptual audio coders is typically performed using a (nonuniform) scalar quantizer (SQ) whose output is entropy-coded, since it allows for simultaneous exploitation of perceptual and statistical source redundancies. Entropy-coded SQ finds a wide range of applications in compression of other source signals as well [17], [18].

A major objection to incorporating bit-rate scalability within existing coders is the resulting loss in performance relative to nonscalable coding. In the most common approach to scalable coding, the “multilayer” approach, the base-layer coder produces the core bit stream and the enhancement layer refines the reconstruction quality by quantizing the base-layer reconstruction error. Conventional scalable systems that minimize a non-mse measure underperform due to two main reasons: 1) each encoding layer fails to fully account for the information conveyed by the preceding layer by operating *only* on the reconstruction error and 2) although the conditional probability density function (pdf) of the source given the base-layer reconstruction differs from the original source pdf, a scaled version of the *same* quantizer is employed for all the encoding layers. In particular, AAC incurs a substantial performance penalty in offering bit-rate scalability when encoding modules operate at low rates.

In this paper, we focus on improving the bit-rate scalability of compression systems that rely on optimization of the WSE distortion metric. We start with a formal proof demonstrating the suboptimality of the conventional approach. Next, we propose two quantization techniques for improving the bit-rate scalability of compression systems. Our first proposed technique—the companded scalable quantization (CSQ) scheme [19], [20]—achieves asymptotically (high resolution) optimal scalability in the operational rate-distortion (RD) sense; preliminary work [19], [20] referred to CSQ as the asymptotically

Manuscript received January 19, 2003; revised March 18, 2005. This work was supported in part by the National Science Foundation under Grants MIP-9707764, EIA-9986057, and EIA-0080134, the University of California MICRO Program, Dolby Laboratories, Inc., Lucent Technologies, Inc., Mind-speed Technologies, and Qualcomm, Inc. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ravi P. Ramachandran.

A. Aggarwal is with Harman Consumer Group, Northridge, CA 91329 USA (e-mail: aaggarwa@harman.com).

S. L. Regunathan is with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: shrane@microsoft.com).

K. Rose is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560, USA (e-mail: rose@ece.ucsb.edu).

Digital Object Identifier 10.1109/TSA.2005.858043

optimal scalable (AOS) scheme. Here, we consider the compandor domain representation of the SQ and note that WSE optimization in the original signal domain is equivalent to mse optimization of the companded signal. Moreover, we observe that it is possible to quantize the reconstruction error without loss of optimality whenever the distortion measure is mse. We hence propose the CSQ approach, which achieves asymptotic optimal scalability by quantizing the reconstruction error in the compandor's *companded* domain. We then fundamentally extend the CSQ scheme by conditioning enhancement-layer quantization on base-layer information [21], [22]. The conditional enhancement-layer quantization (CELQ) scheme is our second proposed technique. In the important case that the source is well modeled by a Laplacian process, we show that CELQ is implementable by switching between two distinct quantizers depending on whether or not the base-layer quantizer employed the so-called "zero dead-zone." CELQ is incorporated in a straightforward manner within the CSQ scheme and shown to further improve its performance by a substantial margin. Simulation results are first given for a synthetic memoryless Laplacian source under a μ -law companded system. The proposed quantization schemes are then implemented within the multilayer MPEG AAC with a standard compatible base layer. Objective and subjective evaluations confirm that major performance gains, by a factor of 2 to 3 in bit-rate savings, can be achieved over standard scalable MPEG AAC. For example, the proposed four-layer scalable coder consisting of 16-kb/s layers achieves performance close to a 64-kb/s nonscalable coder on the standard test database of 44.1-kHz (monophonic) audio. Moreover, the proposed schemes eliminate the computationally demanding parameter optimization procedure at the enhancement layers of the scalable AAC.

The organization of the paper is as follows. Notation and preliminaries are covered in Section II. A brief overview of scalable quantization and audio coding is provided in Section III. The quantizer design problem and a formal proof demonstrating the suboptimality of the conventional approach are given in Section IV. The two proposed schemes, CSQ and CELQ, are detailed in Sections V and VI, respectively. Implementation of the proposed schemes within MPEG AAC is explained in Section VII. Simulation results for synthetic memoryless Laplacian sources with μ -law companding and for real-world audio coding in conjunction with AAC, are presented in Section VIII.

II. NOTATION AND PRELIMINARIES

Throughout the paper we use the following notation. Let the input to the system be random variable X with pdf $f_X(x)$, where an instance of X is denoted by x , and $x \in \mathcal{R}$. Expectation with respect to f_X is denoted by $E_X[\cdot]$, and the differential entropy by $h(X)$.

The nonuniform SQ is denoted by the function $Q(x)$, which partitions the real line into N disjoint cells S_i , $1 \leq i \leq N$. The other parameters pertinent to the quantizer are: the volume of each cell $V(x)$, the high-resolution un-normalized quantization level density $\Lambda(x)$, the scaling factor Δ , and the distortion-rate function δ . Rate is denoted by R and distortion by D . The quantization error is denoted by z . Subscripts e and b denote

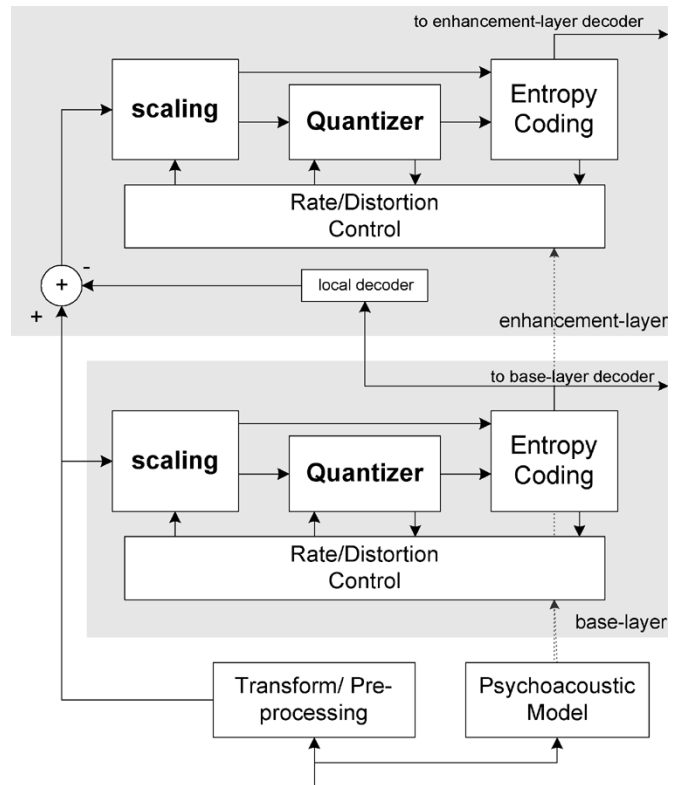


Fig. 1. Block diagram of a two-layer scalable AAC where the enhancement layer quantizes the base-layer reconstruction error. Perceptual redundancy via noise shaping is exploited using the transform and psychoacoustic model. The transform coefficients are grouped into nonuniform bands before adaptive per-band scaling and quantization. Scaling information and quantization indices are transmitted to the decoder for all encoding layers using entropy coding. Target rate is achieved by the rate-distortion control module.

the enhancement-layer and base-layer variables, respectively. A quantized and reconstructed value is denoted by \hat{x} and a predicted value by \tilde{x} , e.g., \hat{x} is the quantized and \tilde{x} is the predicted value of variable x .

III. SCALABLE QUANTIZATION AND AUDIO CODING

In this section, we provide a brief overview of scalable quantization and conventional audio coding techniques. We start with the multilayer AAC since it aptly characterizes a generic perceptual audio coder. For an overview of perceptual audio coding see [23], and for a detailed description of AAC we direct the interested reader to [3]–[6]. Next, we outline relevant well-known results from high-resolution quantization theory. A thorough treatment of the theory is beyond of the scope of this paper and can be found in [24].

A. Scalable Coding of Audio

The most common approach to scalable coding involves the use of multiple encoding layers where each encoding layer quantizes the reconstruction error of the preceding layer. A high-level block diagram of the two-layer AAC is shown in Fig. 1. The time-domain signal is grouped into overlapping frames and transformed into the spectral domain using a modified discrete cosine transform (MDCT). The time-domain data is also input to the psychoacoustic model, whose output is the masking threshold for the spectral coefficients. The transform

coefficients are grouped into nonuniform bands. All coefficients within a given band are quantized using the same nonuniform SQ. The quantization noise is controlled by varying the scaling of a generic quantizer from band to band. Statistical redundancy in the quantized coefficient indices is exploited by the use of entropy coding. The quantization scheme at the enhancement layer is identical to that of the base layer. Quantizer scale-factors that achieve the target distortion are searched for and transmitted at all the encoding layers. The focus of this paper is on the scaling and quantization operations, which are shown in boldface font in the figure.

Other approaches to scalable coding may be grouped into the class of so-called “bit-plane” methods. There is no clear demarcation between multilayer and bit-plane approaches since a multilayer approach may be formulated within the bit-plane scheme and vice-versa. It is, nevertheless, convenient to categorize schemes by their granularity and computational complexity. Bit-plane methods typically offer finer granularity and lower computational complexity than standard multilayer approaches, but their overall RD performance is relatively weaker, as demonstrated in Section VIII. Current bit-plane based scalable coding methods for audio, such as BSAC [25], EZK [26], and ESC [27], use variants of the basic hierarchical partitioning technique, a method originally developed for image coding [28], [29].

B. Entropy-Coded Scalar Quantization

Let $X \in \mathcal{R}$ be a scalar random variable with pdf $f_X(x)$. Quantizer $Q(x)$, maps x into one of N reconstruction points \hat{x}_i , $1 \leq i \leq N$, by partitioning \mathcal{R} into disjoint and exhaustive cells, S_i , $1 \leq i \leq N$, such that $Q(x) = \hat{x}_i$ if $x \in S_i$. We further denote the width of a cell by the function $V(S_i)$. The performance of the quantizer is measured by the WSE distortion criterion given by

$$D = E_X \left[(x - Q(x))^2 w(x) \right] = \int_x (x - Q(x))^2 w(x) f_X(x) dx \quad (1)$$

where $w(x)$ is the weight function.

For large N , high-resolution analysis is employed to model the SQ. It is convenient to define the un-normalized quantization level density function $\Lambda(x)$ [30], where, $\int \Lambda(x) dx = N$, and $1/\Lambda(x)$ gives the width of the quantization cell about x , i.e.,

$$V(S_i) \approx \frac{1}{\Lambda(x)} \quad \text{when } x \in S_i.$$

The following assumptions are typically employed in high-resolution quantization theory when analyzing the RD performance of a quantizer: 1) the source density, $f_X(x)$, is continuous and smooth; 2) the quantization cells are small enough that the source density is nearly constant within the cell; 3) each reconstruction point is located at the center of the cell, i.e., \hat{x}_i is in center of the cell S_i ; and 4) the quantization rate is high, i.e., $N \rightarrow \infty$.

Bennett’s celebrated high-resolution approximation of the quantizer distortion appeared in [31]. It was later extended from mse to the WSE metric, for example in [32]:

$$D(\Lambda) \approx \frac{1}{12} \int_x \frac{w(x) f_X(x)}{\Lambda^2(x)} dx. \quad (2)$$

The quantizer’s rate $R(\Lambda)$ is approximated by the entropy of the quantized output, i.e.,

$$R = - \sum_{i=1}^N p_i \log p_i \quad (3)$$

where, $p_i = P(X \in S_i) = \int_{S_i} f_X(x) dx$.

A high-resolution approximation of R for the quantizer with quantization level density $\Lambda(x)$ was given by Gish and Pierce [33]

$$R(\Lambda) \approx h(X) - E_X \left[\log \left(\frac{1}{\Lambda(x)} \right) \right] \quad (4)$$

where $h(X) = - \int f_X(x) \log(f_X(x)) dx$ is the differential entropy of X . (All logarithms are to the base 2 and the rate is hence measured in bits.)

The asymptotically optimal, nonscalable, entropy-coded SQ is one that minimizes $D(\Lambda)$ subject to the rate constraint $R(\Lambda) \leq R_t$. Let $Q_{\text{ns}}()$ be such an optimal SQ with level density $\Lambda_{\text{ns}}()$. Then

$$\Lambda_{\text{ns}}(x) = \arg \min_{\Lambda(x): R(\Lambda) \leq R_t} D(\Lambda). \quad (5)$$

The optimal high-resolution SQ for the “plain” mse was first outlined by Zador in [34], [35] (see also the rigorous analysis in [30], [33]). Its extension to vector quantization and for the WSE criterion has been proposed by several researchers, e.g., [32]. The optimal, high-resolution SQ is given in terms of the weight function by

$$\Lambda_{\text{ns}}(x) = \frac{\sqrt{w(x)}}{\Delta} \quad (6)$$

where the normalizing constant Δ is chosen to meet the rate constraint. Hence, by (4), we get

$$\log(\Delta) = h(X) - R_t + \frac{1}{2} E_X [\log(w(x))]. \quad (7)$$

The operational distortion-rate function of the nonscalable SQ, denoted $\delta_{\text{ns}}(R)$, is easily derived by substituting (6) and (7) into (2):

$$\delta_{\text{ns}}(R) \triangleq D(R)|_{\Lambda_{\text{ns}}(x)} = \frac{1}{12} 2^{2(h(X)-R)+E_X[\log(w(x))]} \quad (8)$$

IV. SHORTCOMINGS OF CONVENTIONAL SCALABLE CODING

A. Problem and Motivation

Let us focus on the enhancement layer of a multilayer coder. In the conventional scalable system, the enhancement-layer quantization directly quantizes the base-layer reconstruction error. Moreover, the AAC encoder employs a scaled version

of the base-layer quantizer at the enhancement layer. There are two main problems with the conventional approach to scalability.

- 1) The conventional approach misinterprets the perceptual “weights” of the WSE distortion metric at the enhancement layer. At the base layer, a nonuniform quantizer is used to efficiently handle the weights of the distortion metric. Recall that the level density function of the base-layer quantizer was proportional to the square root of the weights [see (6)]. These weights, however, cannot be expressed as a function of the base-layer reconstruction error. Direct quantization of the reconstruction error at the enhancement layer fails to successfully optimize the weighted distortion metric.
- 2) The pdf of the base-layer reconstruction error differs considerably from the source pdf. Restricting the enhancement-layer quantizer to be identical to the base-layer quantizer (up to a scaling factor) is therefore suboptimal. For example, audio signals may be well modeled by the Laplacian density function and the corresponding AAC quantizer may be designed suitably. However, the pdf of the reconstruction error seen by the enhancement layer will be distinctly non-Laplacian. Hence, one quantizer at the enhancement layer (or its scaled version) may not be sufficient to effectively quantize the reconstruction error. In fact, the reconstruction error statistics depend on the quantization level selected at the base layer. An observation that is commonly ignored by conventional approaches.

The effect of poor quantization in scalable AAC is illustrated in Fig. 2. The figure depicts the overall quantizer after two layers of AAC encoding where the base-layer and enhancement-layer scale-factors were chosen arbitrarily. The overall quantizer has numerous obvious artifacts and its resulting operational RD function is poor. Due to poor quantization, the enhancement layer has to search for a new set of quantizer scale-factors and transmit their values as side-information. The side-information representing enhancement-layer scale-factors is, in essence, retransmission of information contained in the base-layer scale-factors. The side-information may be using as much as 30%–40% [36] of the total rate at low encoding rates. It is important to realize that this approach to scalability does not make full use of the available information. In particular, apart from the base-layer reconstruction, the enhancement-layer decoder also has access to the base-layer quantization interval. The base-layer quantizer scale-factors can be used to adjust the scale-factors at the enhancement layer. While it is easy to see that appropriate use of the information from the base layer is the key to better performance of a scalable coder, the means to achieve this goal are not obvious. The benefits of exploiting base-layer information were demonstrated in the context of, scalable predictive coding in [37], multistage vector quantization in [38], and stereo coding of audio in [39].

B. Asymptotic WSE Suboptimality of Conventional Scalability

In this section, we formally prove that, asymptotically, the conventional approach to bit-rate scalability is strictly subop-

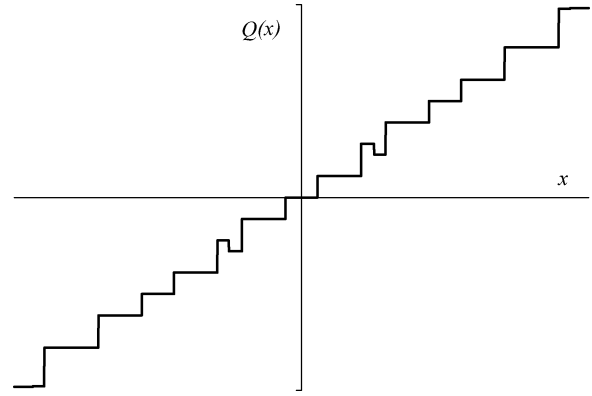


Fig. 2. Overall quantizer after two layers of encoding in AAC. Non-uniform quantization is employed at both the base layer and the enhancement layer. The overall quantizer is poorly designed and remains mismatched to the task of optimizing the distortion metric.

timal for the WSE measure. Consider a two-layer scalable coder. In the standard approach, the enhancement layer simply quantizes the reconstruction error of the base layer. We refer to this scheme as reconstruction error quantization (REQ). Let \hat{x} be the overall reconstructed value of x , z be the reconstruction error at the base layer, and \hat{z} be the quantized value of z . For the base layer we use the terminology described earlier, i.e., \hat{x}_i is the reconstructed value of x when $x \in S_i$, $1 \leq i \leq N_b$. The quantization level density function of the base-layer quantizer is denoted by $\Lambda_b(x)$ and the normalizing constant by Δ_b . For the enhancement layer we use subscript e instead of b and the quantizer input is z , i.e., the enhancement-layer quantizer, $Q_e(z)$, maps $z \in \mathcal{R}$ into one of N_e reconstruction points \hat{z}_j , $1 \leq j \leq N_e$, by partitioning \mathcal{R} into disjoint and exhaustive cells, \tilde{S}_j , $1 \leq j \leq N_e$, such that $Q_e(z) = \hat{z}_j$ if $z \in \tilde{S}_j$. The level density function and the normalizing constant at the enhancement layer are denoted by $\Lambda_e(x)$ and Δ_e , respectively.

In deriving the RD function for the REQ scheme, we make the usual assumptions of high-rate quantization theory: 1) the source density, $f_X(x)$, and the weight, $w(x)$, are continuous and smooth; 2) the quantization cells are small enough that the source density and the weights are nearly constant within the cell; 3) each reconstruction point is located at the center of the cell, i.e., \hat{x}_i is in center of the cell S_i , and \hat{z}_j in center of the cell \tilde{S}_j ; and 4) the quantization rate at each layer is high, i.e., $N_b, N_e \rightarrow \infty$.

Claim 1: Given an optimal base layer, the high-resolution RD performance of a scalable coder employing REQ is strictly worse than the operational distortion-rate function of the non-scalable coder given by (8).

Proof: The overall distortion for the REQ scheme is

$$\begin{aligned}
 D_{\text{req}} &= \int (x - \hat{x})^2 w(x) f_X(x) dx \\
 &\approx \sum_i \int_{x \in S_i} (x - \hat{x})^2 w(\hat{x}_i) f_X(\hat{x}_i) dx \\
 &= \sum_i w(\hat{x}_i) f_X(\hat{x}_i) \int_{-\frac{1}{2\Lambda_b(\hat{x}_i)}}^{\frac{1}{2\Lambda_b(\hat{x}_i)}} (z - \hat{z})^2 dz \quad (9)
 \end{aligned}$$

where the approximation is due to the high-resolution assumptions. The integral in (9) may be evaluated as

$$\begin{aligned} \int (z - \hat{z})^2 dz &= \sum_j \int_{\tilde{S}_j} (z - \hat{z}_j)^2 dz \\ &= \sum_j \int_{-\frac{1}{2\Lambda_e(\hat{z}_j)}}^{\frac{1}{2\Lambda_e(\hat{z}_j)}} z^2 dz \\ &= \sum_i \frac{1}{12\Lambda_e^3(\hat{z}_i)}. \end{aligned}$$

The resulting sum may be viewed as representing a Riemann integral due to the high resolution of the enhancement layer $1/\Lambda_e(\hat{z}_i) = V(\tilde{S}_i) \rightarrow 0$. Hence

$$\int (z - \hat{z})^2 dz \approx \int \frac{1}{12\Lambda_e^2(z)} dz. \quad (10)$$

Substituting (10) in (9), we obtain

$$D_{\text{req}} \approx \frac{1}{12} \sum_i w(\hat{x}_i) f_X(\hat{x}_i) \int_{|z| \leq \frac{1}{2\Lambda_b(\hat{x}_i)}} \frac{1}{\Lambda_e^2(z)} dz.$$

Using a similar Riemann integral argument for the base layer $1/\Lambda_b(x) = V(S_i) \rightarrow 0$ leads to

$$\begin{aligned} D_{\text{req}} &= \frac{1}{12} \int_x w(x) \Lambda_b(x) f_X(x) \int_{|z| \leq \frac{1}{2\Lambda_b(x)}} \frac{1}{\Lambda_e^2(z)} dz dx \\ &= \frac{1}{12} \int_z \frac{K(z)}{\Lambda_e^2(z)} dz \end{aligned} \quad (11)$$

where

$$K(z) = \int_{x: \Lambda_b(x) \leq \frac{1}{2|z|}} w(x) \Lambda_b(x) f_X(x) dx. \quad (12)$$

The base-layer and enhancement-layer rates are related to their respective quantizers by [33]

$$\begin{aligned} R_b &= h(X) + E_X [\log(\Lambda_b(x))] \\ R_e &= h(Z) + E_Z [\log(\Lambda_e(z))]. \end{aligned} \quad (13)$$

To prove that the RD performance of the REQ scheme is strictly worse than that of the non-scalable scheme, we establish two inequalities, one for the distortion and the other for the rate. Using Hölder's inequality [40], we get (14):

$$\begin{aligned} \left[\int_{\Lambda_b(x) \leq \frac{1}{2|z|}} \Lambda_b^3(x) f_X(x) dx \right]^{\frac{1}{3}} &\left[\int_{\Lambda_b(x) \leq \frac{1}{2|z|}} f_X(x) dx \right]^{\frac{2}{3}} \\ &\geq \int_{\Lambda_b(x) \leq \frac{1}{2|z|}} \Lambda_b(x) f_X(x) dx. \end{aligned} \quad (14)$$

Further, since $f_X(x)$ is a pdf, we have

$$\left[\int_{\Lambda_b(x) \leq \frac{1}{2|z|}} f_X(x) dx \right]^{\frac{2}{3}} \leq 1 \quad (15)$$

and, from [41], the pdf of z is given by

$$\begin{aligned} f_Z(z) &= \sum_i f_Z(z|X \in S_i) Pr(X \in S_i) \\ &= \sum_i \begin{cases} f_X(\hat{x}_i), & |z| \leq \frac{1}{2\Lambda_b(\hat{x}_i)} \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

which, as $1/\Lambda_b(\hat{x}_i) \rightarrow 0$, gives

$$f_Z(z) = \int_{x: \Lambda_b(x) \leq \frac{1}{2|z|}} \Lambda_b(x) f_X(x) dx. \quad (16)$$

Substituting (15) and (16) in (14), Hölder's inequality yields

$$\int_{x: \Lambda_b(x) \leq \frac{1}{2|z|}} \Lambda_b^3(x) f_X(x) dx \geq f_Z^3(z). \quad (17)$$

For the REQ to be optimal, the base layer must be optimal in the first place. The optimal base-layer quantizer is one that satisfies $w(x) = \Delta_b^2 \Lambda_b^2(x)$. Therefore, $K(z)$ in (12) reduces to

$$\begin{aligned} K(z) &= \Delta_b^2 \int_{x: \Lambda_b(x) \leq \frac{1}{2|z|}} \Lambda_b^3(x) f_X(x) dx \\ &\geq \Delta_b^2 f_Z^3(z) \end{aligned} \quad (18)$$

where use is made of (17). Now, substituting (18) in (11), we obtain the ‘‘distortion-inequality’’

$$\begin{aligned} D_{\text{req}} &\geq \frac{\Delta_b^2}{12} \int_z \frac{f_Z^2(z)}{\Lambda_e^2(z)} f_Z(z) dz \\ &= \frac{\Delta_b^2}{12} E_Z \left[\frac{f_Z^2(z)}{\Lambda_e^2(z)} \right]. \end{aligned} \quad (19)$$

Next we establish the ‘‘rate-inequality.’’ Using (13), the total rate $R = R_b + R_e$ is given as

$$\begin{aligned} R &= h(X) + E_X [\log(\Lambda_b(x))] + h(Z) + E_Z [\log(\Lambda_e(z))], \\ &= h(X) + E_X [\log(\Lambda_b(x))] - E_Z \left[\log \left(\frac{f_Z(z)}{\Lambda_e(z)} \right) \right] \end{aligned}$$

which implies that

$$E_Z \left[\log \left(\frac{f_Z(z)}{\Lambda_e(z)} \right) \right] = h(X) - R + E_X [\log(\Lambda_b(x))].$$

By applying Jensen's inequality, we obtain

$$\log \left(E_Z \left[\frac{f_Z(z)}{\Lambda_e(z)} \right] \right) \geq h(X) - R + E_X [\log(\Lambda_b(x))]. \quad (20)$$

By combining (19) and (20), and using the fact that $\Lambda_b(x) = \sqrt{w(x)}/\Delta_b$, we get

$$\begin{aligned} D_{\text{req}}(R) &\geq \frac{\Delta_b^2}{12} E_Z \left[\frac{f_Z^2(z)}{\Lambda_e^2(z)} \right] \\ &\geq \frac{1}{12} 2^{2(h(X)-R)+E_X[\log(w(x))]} \end{aligned} \quad (21)$$

Therefore

$$D_{\text{req}}(R_b + R_e) \geq \delta_{\text{ns}}(R_b + R_e). \quad (22)$$

The performance of REQ is strictly worse than the bound unless the two main inequalities, (14) and (20), are satisfied with equality. Inequality (14) is satisfied with equality only if $\Lambda_b^3(x)f_X(x)$ is proportional to $f_X(x)$, hence only if $\Lambda_b(x)$ is a constant, i.e., if the base-layer quantizer is uniform. However, the optimal SQ at the base layer is given by $\Lambda_b(x) = \sqrt{w(x)}/\Delta_b$, and is not uniform for the WSE measure (except when it degenerates to plain mse). Consequently, the RD performance of REQ is *strictly* worse than that of the nonscalable coder for the WSE distortion metric. ■

An interesting observation is that (20) is satisfied with equality only if $\Lambda_e(z)$ is proportional to $f_Z(z)$. Since $\Lambda_e(z)$ is the only parameter available for control at the enhancement layer, the overall distortion of the REQ scheme with an optimum base-layer quantizer is minimized given the total rate when $\Lambda_e(z)$ is proportional to $f_Z(z)$. Hence, the best enhancement-layer quantizer in the REQ approach has level density proportional to the pdf of the reconstruction error.

V. COMPANDED SCALABLE QUANTIZATION

A scalable coder offers scalability at no rate loss if it achieves the RD performance of the nonscalable coder, δ_{ns} . Therefore, δ_{ns} represents the *operational* RD bound for the scalable coder. Further, achieving δ_{ns} yields *asymptotic* optimality because of the high-resolution assumptions employed in deriving this bound. In this section, we develop the CSQ scheme and show that it achieves asymptotic operational RD optimality. One way to obtain optimal scalability is to exploit all the information from the base layer by designing a separate enhancement-layer SQ for each base-layer index. The obvious drawback of such a scheme is its heavy memory requirement. Instead, we consider the compandor-equivalent representation of the (nonuniform) SQ, which consists of a compressor, a uniform SQ and an expander (inverse compressor). Such a mapping is shown in Fig. 3. The compressor function, $c(x)$, and the uniform SQ stepsize, Δ , of the compandor-equivalent representation are related to the quantization level density of the nonuniform SQ, $\Lambda(x)$, by

$$\frac{\partial c(x)}{\partial x} = c'(x) = \Delta \Lambda(x). \quad (23)$$

We can retain the simplicity of an REQ scheme and nevertheless achieve asymptotically optimal performance by taking advantage of the following two observations.

- 1) *REQ is optimal for the mse criterion* ($w(x) = 1$).

For mse and at high resolution, the optimal nonscalable entropy coded SQ is uniform [30], [33], i.e.,

$$w(x) = 1 \Rightarrow \Lambda_{\text{ns}}(x) = \frac{1}{\Delta} \Rightarrow c'(x) = 1. \quad (24)$$

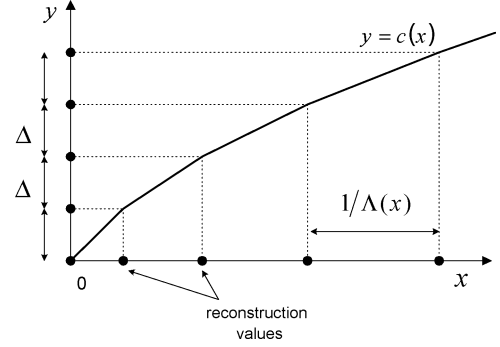


Fig. 3. Block diagram of a nonuniform scalar quantizer and its equivalent companding scheme. The compressor function, $c(x)$, is used as a mapping from a nonuniform to a uniform quantizer. $\Lambda(x)$ is the un-normalized quantization level density of the nonuniform quantizer and Δ is uniform quantizer stepsize.

At high resolution, uniform quantization at the base layer yields uniform pdf of the reconstruction error. This can be formally shown by substituting $\Lambda_b(x) = 1/\Delta_b$ in (16). We get

$$f_z(z) = \begin{cases} \frac{1}{\Delta_b}, & |z| \leq \frac{1}{2\Delta_b} \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

This observation is intuitively obvious—if we employ uniform quantization at the base layer and assume that the source pdf is constant over the quantization interval, the reconstruction error pdf will naturally be uniform. Hence, the optimal SQ at the enhancement layer is also uniform.

To formally prove that REQ is optimal for mse, we simply substitute $\sqrt{w(x)} = 1$, $\Lambda_b(x) = 1/\Delta_b$ and $\Lambda_e(x) = 1/\Delta_e$ in (11), (12) and (13). The resulting set of equations are

$$\begin{aligned} K(z)|_{w(x)=1} &= \int_{x:L(x,z) \leq 1} \Lambda_b(x) f_X(x) dx = f_Z(z) \\ D_{\text{req}}|_{w(x)=1} &= \frac{\Delta_e^2}{12} \int_z K(z) dz = \frac{\Delta_e^2}{12} \\ R_b|_{w(x)=1} &= h(X) - \log(\Delta_b) \\ R_e|_{w(x)=1} &= h(Z) - \log(\Delta_e) \\ &= \log(\Delta_b) - \log(\Delta_e) \\ \Rightarrow D_{\text{req}}|_{w(x)=1} &= \frac{1}{12} 2^{(h(X)-(R_b+R_e))} \\ &= \delta_{\text{ns}}(R_b + R_e)|_{w(x)=1}. \end{aligned}$$

Thus, REQ asymptotically achieves the operational RD bound when the distortion metric is mse. ■

- 2) *For an optimal entropy coded SQ, the WSE of the original signal equals mse of the companded signal.*

The optimal SQ satisfies $\Lambda(x) = \sqrt{w(x)}/\Delta$, which implies that $c'(x) = \sqrt{w(x)}$ reducing Bennett's integral in (2) to $D = \Delta^2/12$. Hence, for the mapping $y = c(x)$, the WSE incurred for quantizing x (in the original domain) equals the mse incurred for quantizing y (in the companded domain). ■

Given the above observations, we construct CSQ as shown by the block diagram of Fig. 4. Let the base-layer and enhancement-layer uniform SQs in the companded domain have

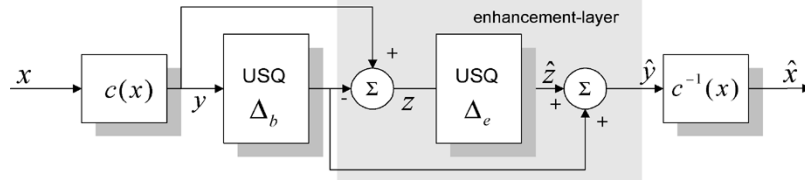


Fig. 4. Block diagram of the CSQ scheme. Input x is first companded using the compressor function $c(x)$. Successive refinement is then performed using uniform scalar quantization (USQ) in the companded domain. Δ_b and Δ_e denote the base-layer and enhancement-layer stepsizes, respectively, and $c^{-1}(x)$ is the expander (inverse compressor).

stepsizes Δ_b and Δ_e , respectively. Let D_{csq} be the distortion of CSQ, and R_b and R_e be the base-layer and enhancement-layer rates. Choose the compressor function such that $c'(x) = \sqrt{w(x)}$ and let $y = c(x)$ be the companded signal.

Claim 2: The RD performance of CSQ achieves the nonscalable RD bound of (8), i.e., CSQ achieves asymptotic optimal scalability.

Proof: The RD performance of the CSQ scheme is obtained as follows:

$$\begin{aligned} D_{csq} &= \frac{\Delta_e^2}{12} \\ R_b &= h(Y) - \log(\Delta_b) \\ &= h(X) + E_X [\log(c'(x))] - \log(\Delta_b) \\ &= h(X) + \frac{1}{2E_X [\log(w(x))]} - \log(\Delta_b) \\ R_e &= \log(\Delta_b) - \log(\Delta_e) \\ \Rightarrow D_{csq} &= \frac{1}{12} 2^{2(h(X) - (R_b + R_e)) + E_X [\log(w(x))]} \\ &= \delta_{ns}(R_b + R_e). \end{aligned}$$

We thus achieve asymptotic optimality. ■

VI. CONDITIONAL ENHANCEMENT-LAYER QUANTIZATION

Let us consider the trade-off between compression efficiency of scalable coding versus its granularity. Given a total target rate, the overall performance of the scalable system improves as fewer layers with larger bit-rate increments are used. At the extreme, the performance of the single layer “scalable” coder is optimal since it is trivially identical to the nonscalable coder. However, scalable quantization may incur heavy performance penalties as the granularity is increased with more layers of smaller bit-rate increments.

In the previous section, scalable coding using CSQ was shown by high-resolution analysis to be asymptotically optimal for the WSE distortion measure. Such analysis assumes that each encoding layer operates at high rates. In this section, we address the more pertinent case where encoding layers operate at low (incremental) rates. The CSQ result derivation employs a compressor function to map WSE in the original signal domain to mse in the companded domain. Successive refinement of the companded signal was then performed using uniform quantizers. An entropy coded uniform quantizer is the optimal SQ for minimizing the mse at high encoding rates. However, the optimal SQ for minimizing mse at low rates may not necessarily be uniform [42], [43]. Approximating the optimal SQ with a uniform SQ may lead to considerable performance degradation at low rates [33]. This observation

is effectively recognized by the designers of practical systems such as AAC, JPEG [44], MPEG-4 video [45] and H.263 [46], which employ an “almost,” but not quite, uniform quantizer in the companded domain. Direct utilization of CSQ may not be possible when the quantizer employed for minimizing the mse metric (in the companded domain) is not uniform. The main problem arises in the compandor-equivalent representation of such an optimal low-rate SQ. The compressor function in this case might neither be smooth nor map the WSE of the original signal to the mse of the companded signal. Hence, the assumptions made in the CSQ derivation are violated and the optimality results are invalid at low encoding rates.

Let us consider the design of the enhancement-layer quantizer when the base layer employs a nonuniform quantizer in the companded domain. Optimality implies achieving the best RD trade-off at the enhancement layer for the given base-layer quantizer. One method to achieve optimality, by brute force, is to design a separate entropy-constrained quantizer for each base-layer reproduction. This approach is prohibitively complex in general. However, in the practically important case where the companded source signal is well modeled by a Laplacian density [47], [48], optimality can be achieved at low complexity by designing different enhancement-layer quantizers for *only* two cases, depending on whether or not the base-layer reproduction is zero. The argument follows from the “memoryless” property of exponential pdfs [49] which can be stated as follows: *given an exponentially distributed variable X consider an interval $[a, b]$, where $0 < a < b$. The conditional pdf of $(X - a)$ is identical for all intervals of the same width $a - b$.* Since the Laplacian is a two sided exponential, the memoryless property extends to the Laplacian pdf for all intervals $[a, b]$ that do not contain the origin.

Let us assume that the source density in the companded domain is modeled by the Laplacian pdf. For a Laplacian pdf, the dead-zone quantizer (DZQ) may be used to closely approximate the optimal entropy coded SQ at all rates [42]. The DZQ has a “dead-zone” around zero whose width is greater than the constant width of all other intervals. Further, the reconstruction levels are shifted toward zero. The structure of the DZQ employed in AAC is shown in Fig. 5. Note that the compressor function to map DZQ to uniform SQ would be nonsmooth.

Claim 3: If a Laplacian source is quantized using DZQ at the base layer, for all nonzero base-layer indices, the reconstructed error is identically distributed (except for obvious possible reversal depending on the sign) and independent of the base-layer reconstruction.

Proof: The proof follows directly from the memoryless property of the exponential pdf. ■

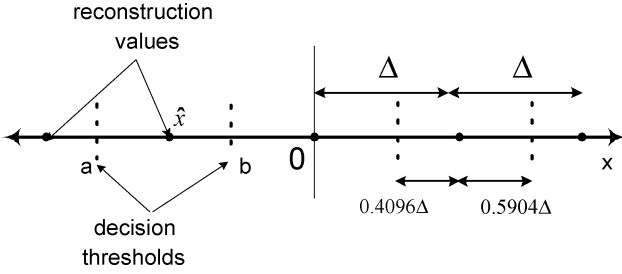


Fig. 5. DZQ used in AAC. The width of quantization interval around zero is nearly 1.2 times the width of other intervals. The reconstruction within an interval is biased toward zero.

This situation is depicted in Fig. 6. Three quantization intervals at the base layer are shown, one where the base-layer reconstruction is zero (region 0), one where it is positive (region 1) and one where it is negative (region 2). It can be clearly seen that whenever the base-layer reconstruction is not zero, the conditional pdf of the signal seen by the enhancement layer is identical in shape (and reversed depending on the sign of base-layer reconstruction). This conditional pdf, $f_X(x/\hat{x}_b)$, is obtained by truncating $f_X(x)$ to the quantization interval $[a, b]$ and normalizing as is shown by the lightly shaded region in the figure. Hence, when the base-layer reconstruction is nonzero, only one quantizer (and its flipped version) is sufficient to optimally quantize the reconstruction error at the enhancement layer. Since speech and audio may be closely approximated by a Laplacian model [47], [48], even in the companded domain, we propose to implement the CELQ scheme by the use of two switchable quantizers. The quantizers are switched depending on whether the base-layer reconstruction is zero or not. Note that the sign of the base-layer reconstruction is available at the enhancement layer eliminating the need to send any side information for flipping of the quantizer.

The design of CELQ can be further simplified, albeit at some loss of optimality. A simple uniform-threshold and reconstruction quantizer (UTRQ) is used at the enhancement layer when the base-layer reconstruction is not zero. A UTRQ is symmetric around zero, has a constant stepsize and, in all the quantization intervals the reproduction value within the interval is “biased” away from the center by a constant number. The only difference between a uniform SQ and UTRQ is the bias away from the center in the latter’s reproduction value. The UTRQ is attractive because of its simple encoding and decoding rule. The encoder simply performs the rounding operation after the coefficient is divided by the stepsize and the decoder subtracts the constant bias from the quantized coefficient index before multiplying it by the stepsize. For the case where the base-layer reconstruction is zero the enhancement layer simply uses a scaled version of the base-layer quantizer.

The proposed CELQ scheme is implemented within the CSQ approach as shown in Fig. 7. The input signal, x , is companded using the compressor function, $c(x)$, to yield the companded signal, y . At the base layer, y is scaled using stepsize Δ_b , and quantized using a fixed DZQ. The reconstruction error, z , is generated in the companded domain and quantized at the enhancement layer. The enhancement-layer switches between DZQ and UTRQ depending on whether or not the base-layer

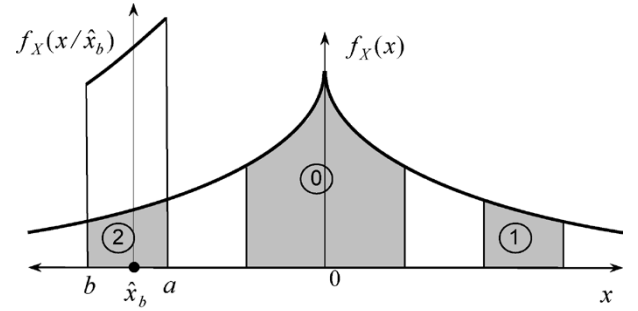


Fig. 6. Shown (shaded area) is the shape of the conditional pdf as seen by the enhancement layer depending on the base-layer quantization interval for three different cases marked as 0, 1, and 2. The actual pdf is obtained by normalizing to one. The conditional pdf when base-layer reconstruction is not zero (cases 1 and 2) is the same except for mirroring.

reconstruction is zero. The enhancement-layer quantizer stepsize is denoted by Δ_e . The overall two-layer quantizer for the CELQ scheme is shown in Fig. 8 (the two-layer quantizer for REQ scheme was shown in Fig. 2). We see from the figure that the overall quantizer exhibits no major artifacts. Furthermore, the width of the quantization intervals increases monotonically away from zero resulting in an improved rate-distortion function, as we demonstrate later.

VII. SCALABLE AAC USING CSQ AND CELQ

In this section, we outline the implementation of the proposed CSQ and CELQ schemes within the scalable AAC. The scaling and quantization operation for the coefficients in one band in AAC is given as

$$i_x = \text{sign}(x) * \text{nint}(\Delta|x|^{0.75} - 0.0946)$$

$$\hat{x} = \text{sign}(i_x) * \left(\frac{|i_x|}{\Delta}\right)^{\frac{4}{3}} \quad (26)$$

where x and \hat{x} are original and quantized coefficients, i_x is the quantized coefficient index, Δ is the quantizer scale-factor, and $\text{nint}()$ and $\text{sign}()$ denote the nearest-integer and signum functions, respectively. From (26), we see that AAC uses a compressor function of $|x|^{0.75}$ and a DZQ to quantize the transform coefficients at the base layer. Further, adaptive scaling per band is applied to shape the quantization noise. Hence, CSQ and CELQ can be implemented in a straightforward manner within AAC.

Let us first focus on the implementation of the CSQ scheme within AAC. At the base layer, once the coefficients are companded and scaled by the appropriate scale-factor, they are *all* quantized using the nearest-integer operation, i.e., the same quantizer. This observation suggests that, if the quantizer scale-factors at the base layer are chosen correctly, optimizing mse in the “companded and scaled domain” is equivalent to optimizing the WSE measure in the original domain. Hence, the enhancement-layer encoder can use a single quantizer in the *companded and scaled* domain for quantizing the reconstruction error. The block diagram of the CSQ scheme when applied to AAC is shown in Fig. 9. The scheme is referred to as CSQ-AAC for brevity. At the base layer, the transform coefficients are grouped into n bands, companded using the compressor function $c(x) = |x|^{0.75}$, scaled using scale-factors

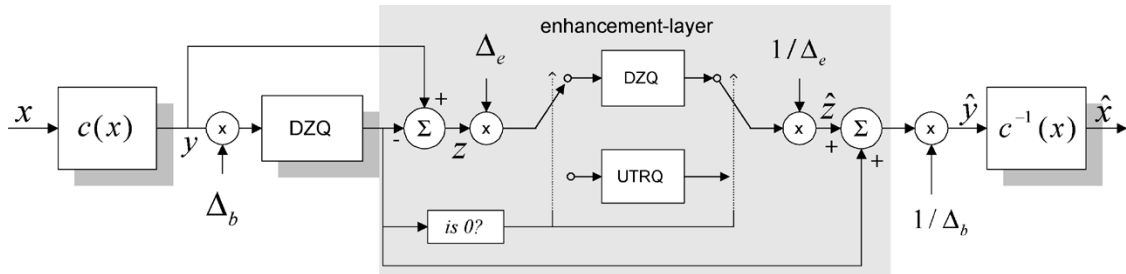


Fig. 7. Block diagram of the CELQ implemented within the companded scalable quantization scheme. In addition to quantizing the reconstruction error in the companded domain, the quantizer at the enhancement layer is switched between the DZQ and the UTRQ depending on whether the base-layer reconstruction is zero or not. $c(x)$ is the compressor function and, Δ_b and Δ_e are the base-layer and enhancement-layer stepsizes, respectively.

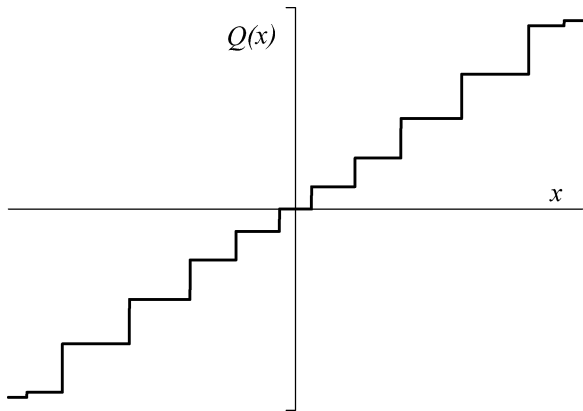


Fig. 8. Overall quantizer after two layers of encoding using CELQ. The overall quantizer is similar to a single-layer (non-scalable) quantizer.

$\Delta_1, \dots, \Delta_n$, and quantized using DZQ. x_1, \dots, x_n are used as representatives for the coefficients in the bands. Note that the base layer is standard compatible. At the enhancement layer, the reconstruction error is generated (by the first adder) in the companded and scaled domain. Coefficients from all the bands are then quantized using a single enhancement-layer quantizer, identical to that of the base layer. The scaling factor for this enhancement-layer quantizer (Δ_e) is chosen in order to meet the bit-rate requirement at that layer. By quantizing the reconstruction error in the scaled and companded domain, the enhancement layer in the CSQ-AAC is able to effectively handle the WSE measure with only one quantizer for all the coefficients. *Major savings in bit rate are achieved since CSQ eliminates the need to transmit the quantizer scale-factor information for each band at the enhancement layers.* In effect, the quantizer scale-factors at the enhancement layer may be viewed as being predicted from those at the base layer. The CSQ scheme has another key benefit over the standard approach. Since quantizer scale-factors are directly available from the base layer, CSQ eliminates the need for their optimization at the enhancement layer. This is significant since parameter search procedures constitute by far the most complex portion of AAC. In fact, barring a small overhead incurred at each layer, the complexity of CSQ is comparable to a single layer AAC.

Since the transform coefficients of a typical audio signal are reasonably modeled by the Laplacian pdf, and AAC uses DZQ at the base layer, CELQ is also implemented within the scalable AAC in a straightforward manner. CELQ may be added to CSQ-AAC by simply putting in a decision block to decide whether or not the quantized value at the base layer is zero. This

choice is made for every coefficient. If the quantized value at the base layer is zero, the enhancement layer continues to use DZQ for quantizing the reconstruction error, otherwise, the enhancement-layer quantizer is switched to use a UTRQ. The reconstruction value of the UTRQ is shifted toward zero by an amount similar to AAC. Further, the reconstruction values of the DZQ and UTRQ are adjusted so that they are always within the base-layer quantization interval.

A block diagram for CELQ-AAC is shown in Fig. 10. It is simplified, without losing representation accuracy, by eliminating the generation of the reconstruction error at the enhancement layer. Instead, the base-layer quantizer directly feeds the quantization interval (in the companded and scaled domain) to the enhancement-layer quantizer, which then refines the quality of the base-layer reconstruction. The advantages of CSQ are retained by operating in the scaled and companded domain. The base layer of CELQ-AAC is also standard compatible.

VIII. SIMULATION RESULTS

Two sets of simulation results are provided. First, we compare the RD performance of the proposed scheme for a synthetically generated memoryless Laplacian source. Then, we evaluate the operation of the proposed schemes on real-world audio signals using AAC.

A. Synthetic Laplacian Source

The performance of the CELQ and CSQ schemes is compared experimentally to the REQ scheme for a two-layer system built using the μ -law companding function ([18, ch. 4]). The input is generated such that the density of the source in the companded domain is (memoryless) Laplacian with zero mean and variance $\sigma^2 = 100$. The maximum value x_m of the μ -law is set to 7σ and hence, the variance of the source has no effect on the operational RD curves. Ideally, the weights of the WSE criterion are defined by the problem at hand. In this synthetic source case, to implement a WSE criterion such that the μ -law is the optimal compressor function, we use a “reverse engineering” approach, i.e., we compute the weights using the relation $\sqrt{w(x)} = c'(x)$. The weights derived for the μ -law compressor function are

$$w(x) = \left(\frac{x_m}{\ln(1 + \mu)} \frac{\frac{\mu}{x_m}}{1 + \frac{\mu x}{x_m}} \right)^2.$$

In our implementation of the REQ scheme, the compressor functions for the base layer and enhancement layer are identical. Further, all the schemes use the same base layer.

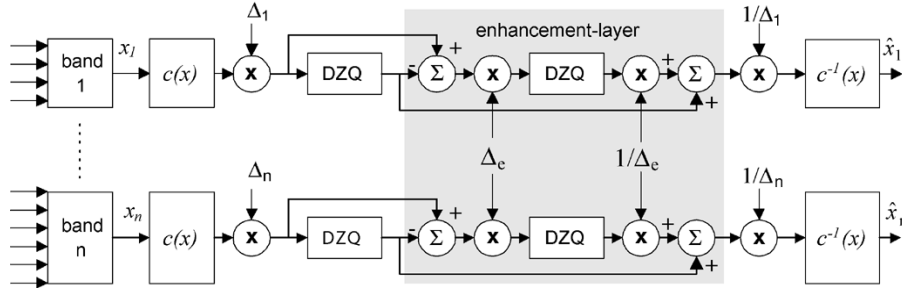


Fig. 9. Block diagram of the companded scalable quantization scheme applied on AAC. The base layer is identical to the standard AAC. x_1, \dots, x_n are used as representatives for the coefficients in the n scale-factor bands. $c(x) = |x|^{0.75}$ denotes the compressor function and $\Delta_1, \dots, \Delta_n$ denotes the base-layer stepsizes for the n bands. The reconstruction error at the enhancement layer is generated in the scaled and companded domain. All coefficients at the enhancement layer are quantized using a single quantizer with stepsize Δ_e .

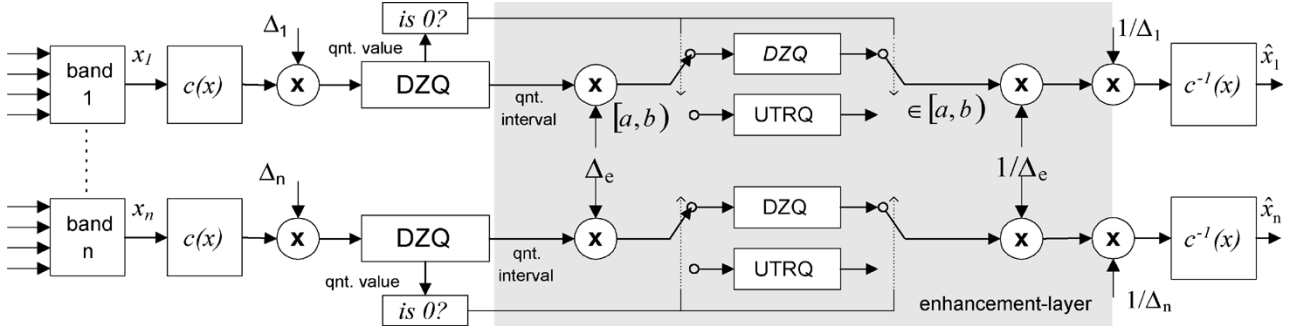


Fig. 10. Block diagram of the conditional enhancement-layer quantizer implemented within the AAC scheme. The base-layer quantization interval in the scaled and companded domain is fed to the enhancement layer. The quantization index is generated at the enhancement layer in a manner consistent with this interval by switching between one of two quantizers, DZQ and UTRQ, depending on whether the base-layer reconstruction is zero or not, respectively.

1) *CSQ Versus REQ: RD Values for Different μ and Base-Layer Rates:* In Fig. 11, we compare the operational RD curves of the CSQ and the REQ schemes for different values of the compressor function parameter and for different base-layer rates. The abscissa denotes the average WSE in decibels and the ordinate gives the total rate in bits/sample. The total rate is the sum of the base-layer and enhancement-layer rates. The CSQ scheme is implemented, as shown in Fig. 4. Uniform SQ is used in the companded domain to quantize the coefficients for both the competing schemes. The figure shows the average RD performance of the competing schemes obtained by connecting the operating points of the scheme. Points on the RD curve are obtained by varying the enhancement-layer quantizer stepsize. Also shown for reference is the RD performance of a single-layer (non-scalable) coder. It is obtained as the convex-hull of *all* the operating points, including the non-scalable coder, and represents the operational RD bound of the coder.

Fig. 11(a) and (b) depicts the behavior of the systems for different values of the compressor function parameter μ . The base-layer rate in these is kept constant at 1 bit/sample and the value of parameter μ is varied from 50 to 255. Fig. 11(b) shows the RD performance of the schemes for a higher value of $\mu = 255$, while the curves in (a) are obtained for the lower value of $\mu = 50$. As expected, we observe that the performance gains for the CSQ over REQ are higher for larger values of μ , i.e., for stronger companding. As we decrease μ , the WSE criterion gradually approaches mse and REQ approaches CSQ (identical for $\mu = 0$).

The effect of different base-layer rates on the RD performance of the schemes is shown via Fig. 11(b) and (c). In these, μ

is kept constant at 255 and the RD performance is evaluated for two different base-layer rates of 1 and 2.5 bits/sample. Fig. 11(c) shows the RD performance of the competing schemes for a (relatively) high base-layer rate of 2.5 bits/sample. When the base-layer and enhancement-layer rates are both high, performance of CSQ nearly equals the non-scalable performance. This can be seen from Fig. 11(c) by comparing the CSQ and non-scalable curves at enhancement-layer rates of ≥ 1.5 bits/sample (total rate ≥ 4 bits/sample). CSQ achieves an asymptotic gain of ≈ 0.75 bits/sample over REQ [indicated by the arrow on Fig. 11(c)]. A limiting (low-rate) case for base-layer rate of 1 bit/sample is shown in (b). We observe that as the base-layer rate decreases the asymptotic performance gain of CSQ over REQ decreases. This can be seen by comparing (b) and (c). CSQ and REQ are identical when the base-layer rate is zero. It is interesting to observe the region where CSQ and REQ curves intersect in (a) and (b). Both the base and enhancement layers operate at the low rate of 1 bit/sample. In this region some of the RD performance of the REQ scheme is better than CSQ. This is due to the fact that a uniform SQ is optimal only asymptotically and CSQ ceases to be efficient when quantization is performed at very low rates.

2) *CELQ Versus CSQ: RD Values for Different Base-Layer Rates:* In Fig. 12, we compare the operational RD curves of CELQ and CSQ at different base-layer rates. The value of μ is kept constant at 255. The key benefit of the CELQ approach is when quantization at the base layer is performed at low rates. Hence, the simulation results are obtained at very low base-layer rates of 0.5 bits/sample and 0.25 bits/sample (Fig. 12(a) and (b), respectively). Unlike the prior set of results, where quantization in the companded domain was performed using a uniform SQ,

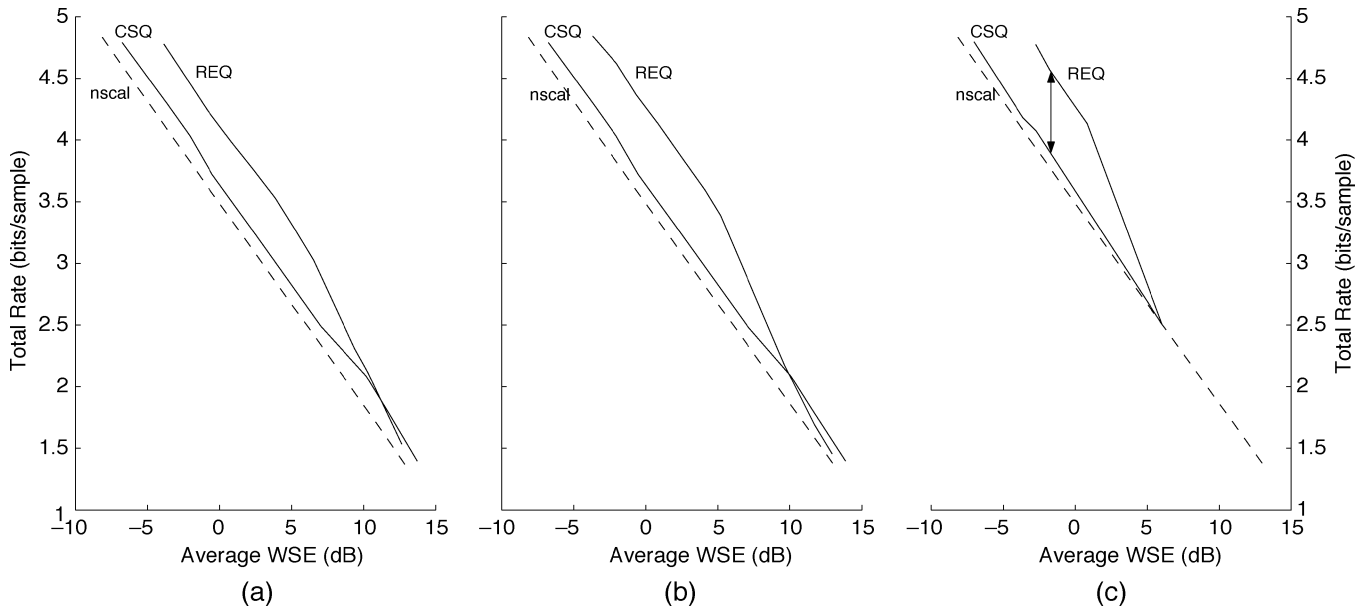


Fig. 11. Performance of a two-layer coder with μ -law companding for the memoryless Laplacian source ($\sigma^2 = 100, x_m = 7\sigma$). The plot shows the total rate (bits/sample) versus WSE (dB) for the two competing methods: CSQ and REQ. The companded signal is quantized using a uniform scalar quantizer. The figure depicts the behavior of the schemes for different degrees of companding and base-layer rates. The performance of a nonscalable coder is shown for reference. (μ , base-layer rate)=(a) (50, 1 bit/sample), (b) (255, 1 bit/sample), and (c) (255, 2.5 bits/sample).

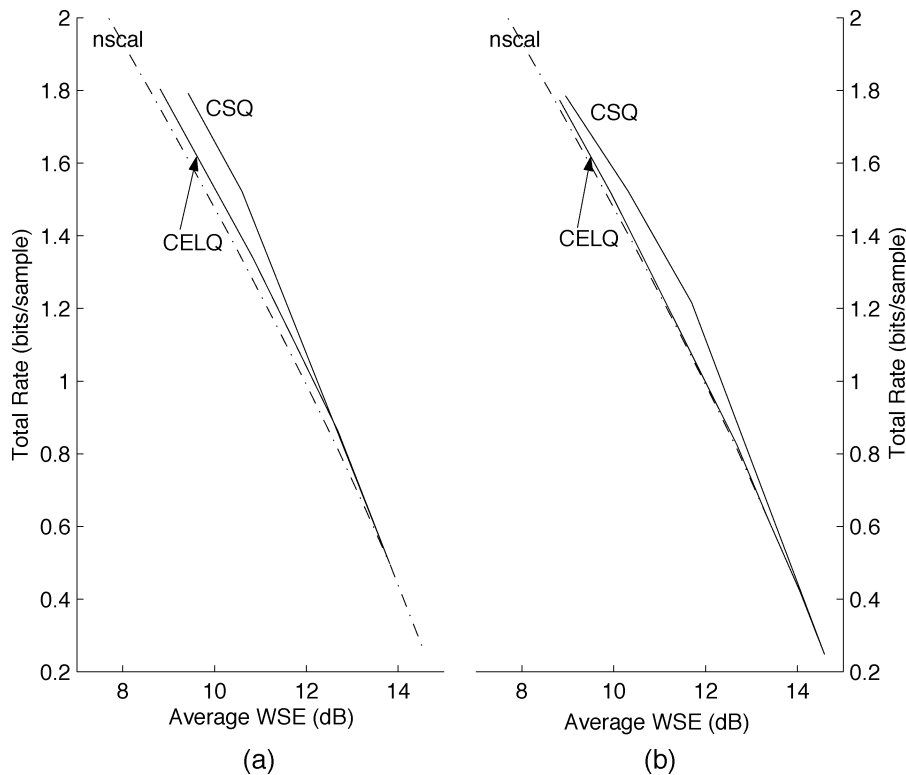


Fig. 12. Performance of a two-layer coder with μ -law companding for the memoryless Laplacian source ($\sigma^2 = 100, x_m = 7\sigma$). The plot shows the total rate (bits/sample) versus WSE (dB) for the two competing methods: CELQ and CSQ. The companded signal is quantized using a dead-zone quantizer. The performance of a nonscalable coder is shown for reference. The figure depicts the behavior of the schemes for different two different base-layer rates: (a) 0.5 bits/sample and (b) 0.25 bits/sample.

a DZQ is employed in the companded domain to quantize the coefficients. The ratio of the quantization interval around zero to other intervals for the DZQ is set to 1.4. For the UTRQ in the CELQ scheme, the reconstruction is “biased” toward zero by 0.2 times the quantizer stepsize.

The performance improvement in the CSQ scheme by the use of DZQ instead of uniform SQ can be seen by comparing the CSQ result in Figs. 11(c) and 12(a). Except for the quantizer in use all the other parameters in these two curves are identical. For more details on the comparison between uniform SQ and DZQ

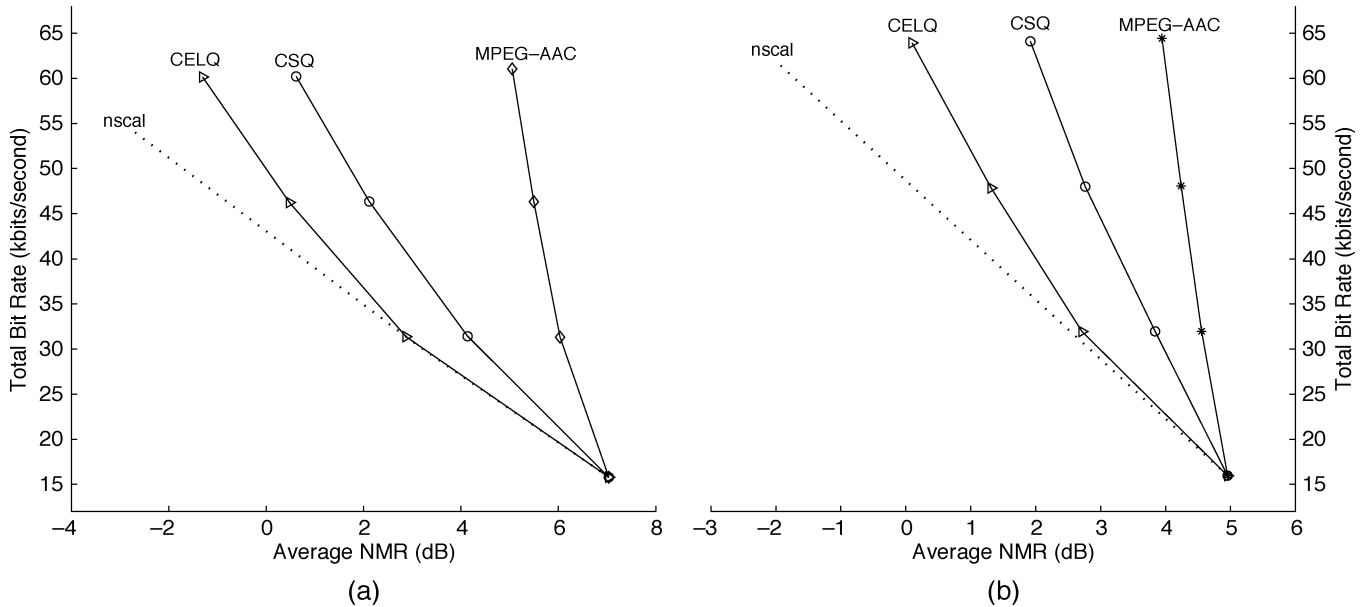


Fig. 13. Performance of a four-layer audio coder with each layer operating at 16 kb/s. Total bit-rate versus Average NMR is shown for the competing methods: CELQ, CSQ and MPEG AAC. Also shown for reference (dotted line) is the performance of a single-layer nonscalable AAC. Averaging is performed over test database of eight critical samples of 44.1-kHz monophonic audio signal. Psychoacoustic model in use is (a) RM-PSY and (b) JJ-PSY.

for the Laplacian source, see [42]. We see that at high rates, the performance of uniform SQ approaches that of DZQ, which is expected since the uniform SQ is asymptotically optimal. However, substantial performance benefit may be achieved at low rates by the use of DZQ.

From Fig. 12 we observe that CELQ adds further improvement to the CSQ scheme and is virtually indistinguishable from the nonscalable curve at all rates. As the base-layer rate is lowered, CSQ approaches the CELQ scheme (identical when base-layer rate is 0).

B. Simulation Results for Scalable MPEG AAC

To demonstrate their practical applicability, in this section, we present objective and subjective test results for the proposed schemes when they are implemented within the scalable MPEG AAC encoder. Implementation details for CSQ and CELQ approaches within the scalable AAC framework were given in Section VII. The MPEG-4 reference model (RM) of AAC [50] is used for simulations and the standard scalable MPEG AAC is built using the REQ approach. For rigorous testing of the proposed schemes the base-line AAC coding module was simplified while keeping standard-compatibility. The bit reservoir was not used and AAC was made to operate at a nearly constant bit-rate. Bandwidth control, window switching and frequency selective switch (FSS) modules were not employed. Design of these tools is beyond the scope of this paper. These simplifications allow for straightforward implementation of the proposed schemes within AAC. To ensure robustness of the proposed approaches for different weighting functions, the schemes are tested with two different psychoacoustic models. The first model is taken from the MPEG-4 RM [50], where the weights in each band are simply a constant factor times the energy in that band. The constant factor depends on the band index. We call this model the RM-PSY. The second model is implemented from [1] and [12] with minor modifications

TABLE I
PERFORMANCE COMPARISON OF THE PROPOSED SCHEMES, CELQ AND CSQ, WITH THE STANDARD APPROACH, REQ, ON A TWO-LAYER SCALABLE AAC CODER. HIGHLIGHTED VALUES SHOW THE RATE SAVINGS OF CELQ OVER REQ FOR SIMILAR DISTORTION VALUES

Total rate (kbps) (base + enhn.)	Average - NMR (dB)		
	CELQ	CSQ	REQ
16 + 16	2.87	4.14	5.22
16 + 32	0.12	1.16	2.04
16 + 48	-2.01	-1.22	-0.96
32 + 32	-1.94	-0.27	0.22
32 + 48	-4.33	-3.21	-2.61
48 + 48	-6.21	-4.54	-4.78

and simplifications. The spreading function and the prediction to find the tonality factor is directly applied to the transform coefficients. We call this model the JJ-PSY. The choice of the psychoacoustic model determines the masking threshold and hence directly affects the resulting NMR calculation. For the test set, eight audio files of sampling rate 44.1 kHz are taken from the EBU-SQAM [50], [51] database, which include tonal signals, castanets, two singing files, and two speech files, one with a male-German speaker. Average NMR is used as the distortion measure in evaluating the objective quality of the companded audio signal. The base layer for all the schemes is identical and standard-compatible.

1) *Objective Results for Two-Layer Coding:* The first set of results are presented for a *two-layer* scalable AAC using the RM-PSY psychoacoustic model. Table I shows the average performance of the competing schemes for the eight test files at different combinations of base-layer and enhancement-layer rates. The results show that CSQ-AAC achieves substantial gains over REQ-AAC in most cases for two-layer scalable coding. CELQ-AAC adds further, and in most cases major, performance benefits over CSQ-AAC. For a clearer picture of the bit-rates savings, we highlight one instance (bold) which

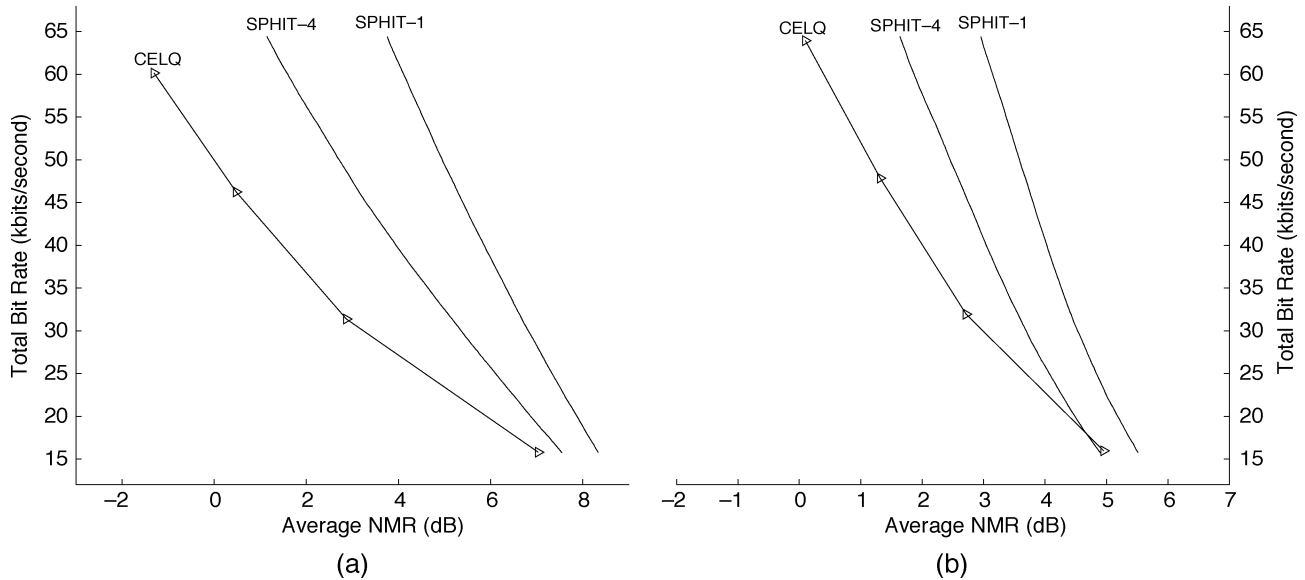


Fig. 14. Performance of a four-layer audio coder with each layer operating at 16 kb/s. The total bit-rate versus Average NMR is shown for the competing methods: CELQ, SPHIT-1 and SPHIT-4. Averaging is performed over a test database of 8 critical samples of 44.1-kHz monophonic audio signal. The psychoacoustic model in use is: (a) RM-PSY and (b) JJ-PSY.

yields comparable distortion. The CELQ coder with 16 + 32 kb/s (base-layer + enhancement-layer rate) is comparable in distortion performance to the REQ scheme with 32 + 32 kb/s, a substantial bit-rate savings of 16 kb/s. The savings are expected to add up with an increasing number of layers, as we demonstrate next.

2) *Objective Results for Multilayer Coding:* In Fig. 13 and 14, we depict the RD curve of a *four-layer* coder with each layer operating at 16 kb/s. In both the figures, the abscissa represents average NMR in decibels and the ordinate gives the total rate in kilobits per second. Average NMR is calculated across the test set of eight audio files. The markers on the curves indicate the RD points at the target rates of 16, 32, 48, and 64 kb/s.

In the first of the two figures, Fig. 13, we compare the performance of the proposed CSQ and CELQ schemes with the standard scalable MPEG AAC which is based on the REQ approach. Fig. 13(a) and (b) shows the performance for the two psychoacoustic modes, RM-PSY and JJ-PSY, respectively. Also shown in the figure by the dotted line is the performance of the non-scalable AAC which serves as the operational RD bound for the scalable coding schemes. From the figure we observe that, for both psychoacoustic models, CSQ gives substantial savings in bit-rate over the standard scalable MPEG AAC for same reproduction quality (as measured by the average NMR). For example, CSQ consumes only 32 kb/s (2×16 kb/s) to achieve an average NMR better than that of the standard scalable MPEG AAC operating at 64 kb/s (4×16 kb/s), a factor of 2 in bit-rate savings. Major performance gains over CSQ are further achieved by the CELQ scheme at virtually no additional computational cost. The CELQ scheme is very close to the non-scalable bound and yields a factor of 3 savings in bit-rate over the standard scalable MPEG AAC. Further, since the quantizer scale-factors at the enhancement layer are directly available from the base layer, the proposed schemes lead to substantial reduction in the search cost, and hence the overall complexity, of scalable AAC when compared with the standard approach.

For comparison with bit-plane-based approaches, we implemented the SPHIT scheme [28], [29], where each parent is associated with a fixed (say n) children clustered together in frequency. We refer to this as SPHIT- n . A simple example of this approach is SPHIT-1, where each parent is associated with only one child—the next transform coefficient in the frame. In [27], SPHIT-4 was shown to outperform EZK [26], which in turn was demonstrated to outperform SPHIT-1. While the authors' proposed method, ESC [27], is claimed to be superior to SPHIT-4, it could not be implemented due to lack of complete algorithmic details in [27]. Hence, SPHIT-4 is apparently the best publicly available bit-plane based scalable approach and we use it for comparison with the proposed scalable approach. In Fig. 14(a) and (b), we compare the CELQ scheme to SPHIT-1 and SPHIT-4. While on the one hand bit-plane based methods offer lower complexity and finer granularity, we see from the figure that the overall RD performance of these scheme is inferior to CELQ.

C. Subjective Results for Multilayer Coding

We performed an informal subjective "AB" comparison test for the CELQ approach consisting of four layers of 16 kb/s each and the non-scalable coder operating at 64 kb/s. The test set contained eight music and speech files from the SQAM database, including castanets and German male speech. A mixed group of eight trained and novice listeners performed the evaluation. Each file was compressed by both competing schemes and the two compressed files were presented in random order to the listener. The listeners were asked to indicate their preference between the two samples and were also provided with the option of choosing "no preference" if no discernible difference was perceived. Table II gives the test results. On an average, listeners rated the overall quality of the CELQ as equivalent to the non-scalable coder. The four-layer scalable CELQ-AAC coder consisting of 16-kb/s layers achieves performance equal to a 64-kb/s non-scalable coder.

TABLE II
SUBJECTIVE PERFORMANCE OF A FOUR-LAYER CELQ (16 × 4 kb/s), AND
NONSCALABLE (64 kb/s) CODER

preferred n-scal @64 kbps	preferred CELQ @16x4 kbps	no preference
26.56%	26.56%	46.88%

As seen from Table I, Fig. 13, Fig. 14, and Table II, CSQ and CELQ achieve major performance gains over the standard scalable AAC and conventional bit-plane based scalable approaches. In particular, CSQ achieves a factor of 2, and CELQ a factor of 3, in bit-rate savings over the standard scalable AAC. Furthermore, both CELQ and CSQ offer reduction in computational complexity over the scalable AAC.

IX. CONCLUSION

In this paper, we presented two quantization techniques for improving bit-rate scalability of compression systems which optimize a weighted squared-error distortion metric and demonstrated their implementation within the multilayer AAC. By operating only on the base-layer reconstruction error, the conventional approach fails to utilize all the information available at the enhancement layer, and the enhancement-layer quantizer remains mismatched to the task of optimizing the distortion measure, resulting in poor scalability. We derived the companded scalable quantization scheme and proved that it achieves asymptotic optimal scalability by quantizing the reconstruction error in the companded domain. It was then extended to the case of low-rate coding by the use of enhancement-layer quantization which is conditional on the base-layer information. It was shown that in the case of Laplacian sources only two switched quantizers at the enhancement layer are needed to approach the RD bound when the base layer employs an optimal entropy-constrained scalar quantizer. Simulation results on the μ -law companding scheme with a memoryless Laplacian source, and on real-world signals using the multilayer AAC, demonstrated that the proposed schemes can achieve significant gains over conventional scalable coding, and eliminate a large fraction of the performance penalty typically incurred by the conventional approach to scalability. Furthermore, the proposed schemes lead to substantial savings in computation complexity for the scalable AAC by eliminating the need to optimize the quantization parameters at the enhancement layer. Informal listening tests indicated that our implementation of a four-layer scalable coder consisting of 16-kb/s layers was shown to achieve performance close to a 64-kb/s nonscalable coder on the standard test database of 44.1-kHz audio.

REFERENCES

- [1] *Information Technology—Very Low Bitrate Audio-Visual Coding*, ISO/IEC Std. ISO/IEC JTC1/SC29 14 496-3:2001(E)—Part 3: Audio, 2001.
- [2] B. Grill, "A bit rate scalable perceptual coder for MPEG-4 audio," in *Proc. 103rd AES Conv.*, New York, 1997, preprint 4620.
- [3] *Information Technology—Generic Coding of Moving Pictures and Associated Audio*, ISO/IEC Std. ISO/IEC JTC1/SC29 13818-7:1997(E)—Part 7: Advanced Audio Coding, 1997.
- [4] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, and Y. Oikawa, "ISO/IEC MPEG-2 advanced audio coding," *J. Audio Eng. Soc.*, vol. 45, no. 10, pp. 789–814, October 1997.
- [5] P. Noll, "MPEG digital audio coding standards," in *Digital Signal Processing Handbook*, V. Madiseti and D. B. Williams, Eds. New York: CRC/IEEE Press, 1998.
- [6] J. Johnston, S. Quackenbush, G. Davidson, and K. Brandenburg, "MPEG audio coding," in *Wavelet, Subband and Block Transforms in Communications and Multimedia*, A. N. Akansu and M. Medley, Eds. Norwell, MA: Kluwer, 1999.
- [7] L. D. Fielder, M. Bosi, G. Davidson, M. Davis, C. Todd, and S. Vernon, "AC-2 and AC-3: low-complexity transform-based audio coding," in *Collected Papers on Digital Audio Bit-Rate Reduction*, N. Gilchrist and C. Grewin, Eds: Audio Engineering Society, 1996, pp. 54–72.
- [8] D. Sinha, J. D. Johnston, S. Dorward, and S. Quackenbush, "The perceptual audio coder (PAC)," in *Digital Signal Processing Handbook*, V. Madiseti and D. B. Williams, Eds. New York: CRC/IEEE Press, 1998.
- [9] K. Akagiri, M. Katakura, H. Yamauchi, E. Saito, M. Kohut, M. Nishiguchi, and K. Tsutsui, "Sony systems," in *Digital Signal Processing Handbook*, V. Madiseti and D. B. Williams, Eds. New York: CRC/IEEE Press, 1998.
- [10] H. Fletcher, "Auditory patterns," *Rev. Modern Phys.*, vol. 12, pp. 47–65, Jan. 1940.
- [11] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, 2nd ed. New York: Springer-Verlag, 1999.
- [12] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Select. Areas Commun.*, vol. 6, no. 2, pp. 314–323, Feb. 1988.
- [13] K. Brandenburg, "Evaluation of quality for audio encoding at low bit rates," in *Proc. 82nd AES Conv.*, 1987, preprint 2433.
- [14] R. J. Beaton, J. G. Beerends, M. Keyhl, and W. C. Treumiet, "Objective perceptual measurement of audio quality," in *Collected Papers on Digital Audio Bit-Rate Reduction*, N. Gilchrist and C. Grewin, Eds: Audio Engineering Society, 1996, pp. 126–152.
- [15] J. G. Beerends and J. A. Stemerink, "A perceptual audio quality measure based on a psychoacoustic sound representation," *J. Audio Eng. Soc.*, vol. 40, no. 12, pp. 963–978, Dec. 1992.
- [16] W. C. Treumiet and G. A. Soudre, "Evaluation of the ITU-R objective audio quality measurement method," *J. Audio Eng. Soc.*, vol. 48, no. 3, pp. 164–173, Mar. 2000.
- [17] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.
- [18] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [19] A. Aggarwal, S. L. Regunathan, and K. Rose, "Asymptotically optimal scalable coding for minimum weighted mean square error," in *Proc. IEEE Data Compression Conf.*, Mar. 2001, pp. 43–52.
- [20] —, "Compander domain approach to scalable AAC," in *Proc. 110th AES Conv.*, Amsterdam, The Netherlands, 2001, preprint 5296.
- [21] A. Aggarwal and K. Rose, "A conditional enhancement-layer quantizer for the scalable MPEG advanced audio coder," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 2, May 2002, pp. 1833–1836.
- [22] —, "Approaches to improve quantization performance over the advanced audio coder," in *Proc. 112th AES Convention*, 2002, preprint 5557.
- [23] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–515, Apr. 2000.
- [24] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [25] S. Park, Y. Kim, and Y. Seo, "Multi-layer bit-sliced bit-rate scalable audio coding," in *Proc. 103rd AES Conv.*, New York, 1997, preprint 4520.
- [26] B. Leslie, C. Dunn, and M. Sandler, "Developments with a zerotree audio codes," in *Proc. AES 17th Int. Conf.*, Sep. 1999, pp. 251–257.
- [27] C. Dunn, "Efficient audio coding with fine-grain scalability," in *Proc. 111th AES Conv.*, New York, 2001, preprint 5492.
- [28] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.
- [29] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, 1996.
- [30] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 4, pp. 373–380, Jul. 1979.

- [31] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, Jul. 1948.
- [32] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1082–1091, May 1999.
- [33] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, no. 5, pp. 676–683, Sep. 1968.
- [34] P. L. Zador, Asymptotic quantization of continuous random variables, in Bell Laboratories Memorandum, unpublished, 1966.
- [35] —, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, pt. 1, vol. IT-28, no. 2, pp. 139–149, Mar. 1982.
- [36] A. Aggarwal, S. L. Regunathan, and K. Rose, "Trellis-based optimization of MPEG-4 advanced audio coding," in *Proc. IEEE Workshop on Speech Coding*, Sep. 2000, pp. 142–144.
- [37] K. Rose and S. L. Regunathan, "Toward optimal scalability in predictive coding," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 965–976, Jul. 2001.
- [38] D. H. Lee, D. L. Neuhoff, and K. K. Paliwal, "Cell-conditioned multistage vector quantization," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 1, 1991, pp. 653–656.
- [39] A. Aggarwal, S. L. Regunathan, and K. Rose, "Optimal prediction in scalable coding of stereophonic audio," in *Proc. 109th AES Conv.*, Los Angeles, CA, 2000, preprint 5273.
- [40] R. M. Gray, *Source Coding Theory*. Norwell, MA: Kluwer, 1990.
- [41] D. H. Lee and D. L. Neuhoff, "Asymptotic distribution of the errors in scalar and vector quantizers," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 446–460, Mar. 1996.
- [42] G. J. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1365–1374, Sep. 1996.
- [43] T. Berger, "Minimum entropy quantizers and permutation codes," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 2, pp. 149–157, Mar. 1982.
- [44] *Information Technology—JPEG 2000 Image Coding System*, ISO/IEC 15444-1:2000, Part 1: Core coding system, 2000.
- [45] *Information Technology—Very Low Bitrate Audio-Visual Coding*, ISO/IEC 14496-2::2001(E), ISO/IEC JTC1/SC29, Part 2: Visual, 2001.
- [46] *Video Coding for Low Bit Rate Communication*, ITU-T Recommendation H.263 Std. ITU-T Recommendation H.263, 1998.
- [47] M. D. Paez and T. H. Glisson, "Minimum mean-squared-error quantization in speech PCM and DPCM systems," *IEEE Trans. Commun.*, vol. COM-20, no. 2, pp. 225–230, Apr. 1972.
- [48] R. C. Reininger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM-31, no. 6, pp. 835–839, Jun. 1983.
- [49] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
- [50] MPEG Audio Web Page.. [Online] Available: <http://www.tnt.uni-hannover.de/project/mpeg/audio/>
- [51] European Broadcasting Union (EBU) Std., "Sound Quality Assessment Material Recordings for Subjective Tests," Rev. Tech. 3253-E, [Online] Available: http://www.ebu.ch/tech32/tech_0253.pdf, Apr. 1988.



Ashish Aggarwal received the B.E degree in electronics from the Bombay University, Bombay, India, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Santa Barbara, in 1998 and 2002, respectively.

From July 2002 to July 2003, he was with PortalPlayer, Inc., Kirkland, WA, where he carried out the design and implementation of audio coders such as MP3 and AAC. In July 2003, he joined Harman International's Advanced Technology Group. His main research activities are audio compression and post-

processing algorithms.

Dr. Aggarwal currently serves as a member of the IEEE Technical Committee on Audio and Electroacoustics. He is a member of the IEEE Signal Processing and Communications Societies and is a member of the AES.

Shankar L. Regunathan received the B.Tech. degree in electronics and communication from the Indian Institute of Technology, Madras, India, in 1994, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Santa Barbara, in 1996 and 2001, respectively.

Currently, he is with Microsoft Corporation, Redmond, WA.



Kenneth Rose (S'85–M'91–SM'01–F'03) received the Ph.D. degree in 1991 from the California Institute of Technology, Pasadena.

In 1991, he joined the Department of Electrical and Computer Engineering, University of California at Santa Barbara, where he is currently a Professor. His main research activities are in information theory, source and channel coding, video and audio coding and networking, pattern recognition, and nonconvex optimization in general. He is also particularly interested in the relations between information theory, estimation theory, and statistical physics, and their potential impact on fundamental and practical problems in diverse disciplines. His optimization algorithms have been adopted and extended by others in numerous disciplines beside electrical engineering and computer science.

Dr. Rose currently serves as an Area Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He co-chaired the Technical Program Committee of the 2001 IEEE Workshop on Multimedia Signal Processing. He was co-recipient of the 1990 William R. Bennett Prize-Paper Award of the IEEE Communications Society, and of the 2004 IEEE Signal Processing Society Best Paper Award (in the area of image and multidimensional signal processing).