Analysis of High Rate LPC Vector Quantizers Designed by Minimizing Suboptimal Error Measures

William R. Gardner[1,2*]

Bhaskar D. Rao[1]

[1]Department of ECE
University of California, San Diego
San Diego, CA 92093-0407

[2]QUALCOMM Inc.
6455 Lusk Blvd.
San Diego, CA 92121

Abstract

For the quantization of the linear predictive coding (LPC) parameters in speech coding systems, the Log Spectral Distortion (LSD) measure is often cited as the performance measure most correlated with speech quality. However, most practical quantization schemes use simpler error measures, such as mean squared error (MSE) or weighted mean squared error (WMSE) measures between the quantized and unquantized LPC coefficients, reflection coefficients, arcsine coefficients, log area ratios, or line spectral pair frequencies (LSPs). This paper develops analytical expressions for the performance of high rate vector quantization (VQ) schemes which are trained by minimizing suboptimal distortion measures, and applies these results to the problem of quantizing the LPC parameters. In particular, the theory is developed to evaluate the performance, as measured by one distortion measure, of a vector quantizer which has been trained by minimizing a different distortion measure. Using this analysis, the performance, in LSD, of vector quantizers trained by minimizing MSE and WMSE measures is theoretically evaluated.

Introduction

Quantization of random vectors has received much attention in recent years due to the ability of vector quantizers to achieve performance superior to that achieved by scalar quantizers [1]. In particular, in many speech coding systems vector quantizers are used to quantize the coefficients of the linear prediction filter (e.g. [2, 3]). Previous work has shown that the Log Spectral Distortion measure is well correlated with the resulting speech quality, and this measure is often used in the LPC-VQ literature as the measure of distortion for performance evaluation [4, 5]. However, the LSD measure is not used for training and quantization in most practical schemes due to the lack of a closed form expression for the generalized centroid, which is needed in most VQ training algorithms, and due to the high complexity required in evaluating the LSD measure during quantization. In most practical schemes, simpler error measures are used. Typically, either MSE or WMSE measures are used on the LPC coefficients, the reflection coefficients, the log area ratios, the arcsine coefficients, or the LSP frequencies. Very little theoretical work has been done on quantifying the effect of using these error measures on the overall performance of the resulting vector quantizers, and most work on these schemes is based primarily on heuristic arguments and experimental results [2, 3]. The overall performance of these different schemes thus far has only been determined by experimentally testing the schemes [5], which gives little information on how to improve the schemes and on how far any particular scheme is from the optimal scheme.

This paper presents a theoretical framework for evaluating the performance, in terms of one distortion measure, of a high rate vector quantizer which has been trained by minimizing a different distortion measure. It is shown that a high rate VQ trained by minimizing a quadratically weighted error measure will converge to a VQ trained by minimizing the original measure, if the quadratic weighting matrix is the "sensitivity matrix," which is the second term of the Taylor series expansion of the original measure. The results presented can be used to determine the performance of MSE and WMSE trained vector quantizers, and to determine the optimal set of WMSE weightings. The theory is then applied to the problem of quantizing LPC coefficients in speech coding systems.

High Rate Vector Quantization

The theoretical performance, as measured by simple distortion measures, of high rate vector quantizers for a random source with a given density function has been investigated by several previous authors [6, 7]. In this paper, the work in [6,7] is extended in two ways so that it may be used to evaluate the performance of suboptimal LPC-VQ schemes. First, the results are extended to a broader class of error measures which may be functions of the absolute location of the input vector, such as the LSD measure. Second, the work is extended to give the high rate performance, measured by distortion measure d_1, of a vector quantizer which was trained by minimizing a different distortion measure, d_2.

Let $\boldsymbol{x}$ be an n dimensional random vector in $\mathcal{R}^n$ with

*This material is based upon work supported under a National Science Foundation Graduate Fellowship.

probability density function $f(\boldsymbol{x}) = f(x_1, x_2, ..., x_n)$. A B bit vector quantizer is composed of a quantization function $Q(\boldsymbol{x})$ which maps $\boldsymbol{x}$ to a codevector $\bar{\boldsymbol{x}}$ in a codebook of 2^B codevectors, $\bar{\boldsymbol{X}} = \{\bar{\boldsymbol{x}}_1, ..., \bar{\boldsymbol{x}}_{2^B}\}$. Let S_i be defined by

$$S_i = \{\boldsymbol{x} | Q(\boldsymbol{x}) = \bar{\boldsymbol{x}}_i\},$$

i.e. S_i is the set of all points quantized to $\bar{\boldsymbol{x}}_i$, or the "Voronoi region" for $\bar{\boldsymbol{x}}_i$. The goal of a quantization scheme is to minimize the expected distortion introduced by quantization, as measured by some error measure E defined by

$$E_d(Q) = \mathcal{E}_{\boldsymbol{x}}(d(\boldsymbol{x}, Q(\boldsymbol{x}))). \quad (1)$$

Here, $\mathcal{E}$ denotes the expectation operator and $d(\boldsymbol{x}, \boldsymbol{y})$ is a distortion function with continuous derivatives. Although many traditional norm based measures are not differentiable at $\boldsymbol{x} = \boldsymbol{y}$, many popular squared error (or 2-norm-squared, as opposed to 2-norm) based measures are continuously differentiable everywhere.

Equation 1 can be expressed as

$$E_d(Q) = \int_{\boldsymbol{x}} d(\boldsymbol{x}, Q(\boldsymbol{x})) f(\boldsymbol{x}) d\boldsymbol{x}. \quad (2)$$

Since $Q(\boldsymbol{x}) = \bar{\boldsymbol{x}}_i; \ \forall \boldsymbol{x} \in S_i$, equation 2 can be rewritten as

$$E_d(Q) = \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} \int_{\boldsymbol{x} \in S_i} d(\boldsymbol{x}, \bar{\boldsymbol{x}}_i) f(\boldsymbol{x}) d\boldsymbol{x}.$$

Previous authors [6,7] have noted that as B gets large the volume of the Voronoi regions gets small, and $f(\boldsymbol{x}) \vec{\approx} f(\bar{\boldsymbol{x}}_i); \ \forall \boldsymbol{x} \in S_i$, if $f(\boldsymbol{x})$ is sufficiently smooth. Hence,

$$E_d(Q) \vec{\approx} \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} f(\bar{\boldsymbol{x}}_i) \int_{\boldsymbol{x} \in S_i} d(\boldsymbol{x}, \bar{\boldsymbol{x}}_i) d\boldsymbol{x}. \quad (3)$$

In this paper, $\vec{\approx}$ denotes an approximation for which the limiting approximation error is zero as the rate of the quantizer approaches infinity. $\approx$ denotes a strict approximation which may have a non-zero approximation error as the rate of the quantizer approaches infinity.

Performing a Taylor series expansion of $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$ about $\boldsymbol{x} = \bar{\boldsymbol{x}}$, holding the second vector in $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$ constant, and noting that the constant and first order terms in the expansion are zero for distortion measures [8–10], results in

$$d(\boldsymbol{x}, \bar{\boldsymbol{x}}) = \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^T \boldsymbol{D}(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) + O(||\boldsymbol{x} - \bar{\boldsymbol{x}}||^3), \quad (4)$$

where $\boldsymbol{D}(\bar{\boldsymbol{x}})$ is an n by n dimensional matrix with j, kth element defined by

$$D_{j,k}(\bar{\boldsymbol{x}}) = \left. \frac{\partial^2 d(\boldsymbol{x}, \bar{\boldsymbol{x}})}{\partial x_j \partial x_k} \right|_{\boldsymbol{x} = \bar{\boldsymbol{x}}}.$$

As B gets large, the density of points in $\bar{\boldsymbol{X}}$ near any particular vector $\boldsymbol{x}$ becomes large [6,7]. Hence $||\boldsymbol{x} - \bar{\boldsymbol{x}}_i||$ gets small $\forall \boldsymbol{x} \in S_i$, so

$$d(\boldsymbol{x}, \bar{\boldsymbol{x}}_i) \vec{\approx} \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}}_i)^T \boldsymbol{D}(\bar{\boldsymbol{x}}_i)(\boldsymbol{x} - \bar{\boldsymbol{x}}_i), \qquad \forall \boldsymbol{x} \in S_i.$$

$\boldsymbol{D}(\boldsymbol{x})$ is denoted here as the "sensitivity matrix." The diagonal elements of the sensitivity matrix are related to the sensitivity of the individual vector elements to quantization error, and the off-diagonal elements are related to the interactions which occur when multiple vector elements are quantized simultaneously. In general, the sensitivity matrix is not diagonal, and a weighted sum of the squared errors of the individual vector elements will not approach the original error measure.

Substituting this result into equation 3,

$$E_d(Q) = \frac{1}{2} \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} f(\bar{\boldsymbol{x}}_i) \int_{\boldsymbol{x}: \boldsymbol{x} + \bar{\boldsymbol{x}}_i \in S_i} \boldsymbol{x}^T \boldsymbol{D}(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x}.$$

If the volume of the Voronoi region S_i is denoted by $vol(S_i)$, then the probability of a vector $\boldsymbol{x}$ falling in that Voronoi region is

$$P_i = Prob(\boldsymbol{x} \in S_i) \vec{\approx} f(\bar{\boldsymbol{x}}_i) vol(S_i), \quad (5)$$

and thus

$$E_d(Q) \vec{\approx} \frac{1}{2} \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} \frac{P_i}{vol(S_i)} \int_{\boldsymbol{x}: \boldsymbol{x} + \bar{\boldsymbol{x}}_i \in S_i} \boldsymbol{x}^T \boldsymbol{D}(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x}.$$

The following lemma allows for simplification of the above expression.

Lemma 1

For any bounded set S_i, $\boldsymbol{M} > 0$,

$$\int_{\boldsymbol{x}: \boldsymbol{x} \in S_i} \boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} d\boldsymbol{x} \geq \int_{\boldsymbol{x} \in T(\boldsymbol{M}, vol(S_i))} \boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} d\boldsymbol{x},$$

where $T(\boldsymbol{M}, V)$ is the "$\boldsymbol{M}$-shaped" hyper-ellipsoidal region with volume V defined by

$$T(\boldsymbol{M}, V) = \left\{ \boldsymbol{x} \left| \left(\frac{\kappa_n^2}{V^2 |\boldsymbol{M}|} \right)^{1/n} \boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} \leq 1 \right. \right\},$$

and where κ_n is the volume of the n-dimensional unit sphere.

The proof of this Lemma exactly follows the proof of Lemma 5.3.1 given in [11]. By using $\boldsymbol{M} = \boldsymbol{D}(\bar{\boldsymbol{x}}_i)$, Lemma 1 states that the distortion incurred by quantizing all points in S_i to $\bar{\boldsymbol{x}}_i$ is greater than the distortion incurred when S_i is replaced by the "$\boldsymbol{D}(\bar{\boldsymbol{x}}_i)$-shaped" hyper-ellipsoid having the volume of S_i. Although Lemma 1 takes the form of a lower bound, it can also be viewed as an approximation to the distortion incurred by quantizing the points in S_i to $\bar{\boldsymbol{x}}_i$, if S_i takes the optimal, lattice shape. Although the hyper-ellipsoids cannot be formed into a lattice as required by the quantizer, the shapes of the Voronoi regions will approach the relative dimensions of the hyper-ellipsoidal regions [9]. This approximation to the shape of the Voronoi regions is a strict approximation. In [12], the error in this approximation was investigated for spaces up to dimension 10, and shown to be below 4 percent for dimensions under 5.

Continuing, the optimal Voronoi regions will have the approximate shapes of $T(\boldsymbol{D}(\bar{\boldsymbol{x}}_i), vol(S_i))$ and thus

$$E_d(Q_{opt}) \approx \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} \frac{P_i}{2vol(S_i)} \int_{\boldsymbol{x} \in T(\boldsymbol{D}(\bar{\boldsymbol{x}}_i), vol(S_i))} \boldsymbol{x}^T \boldsymbol{D}(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x}. \quad (6)$$

In [8,9], is was shown that

$$\int_{\boldsymbol{x} \in T(\boldsymbol{M}, V)} \boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} d\boldsymbol{x} = V \frac{n}{n+2} \left(\frac{V^2}{\kappa_n^2} |\boldsymbol{M}| \right)^{1/n}.$$

Substituting this result into equation 6 gives

$$E_d(Q_{opt}) \approx \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} \frac{nP_i}{2(n+2)} \left(\frac{vol(S_i)^2}{\kappa_n^2} |\boldsymbol{D}(\bar{\boldsymbol{x}}_i)| \right)^{1/n}.$$

This sum can be approximated by an integral, as motivated in [6]. Let $g_B(\boldsymbol{x}) = 1/(2^B vol(S_i))$ and let $\lambda(\boldsymbol{x})$ be the limiting value of $g_B(\boldsymbol{x})$ as $B \to \infty$. $\lambda(\boldsymbol{x})$ has unit integral, and is the limiting density of quantization vectors in $\bar{\boldsymbol{X}}$ near $\boldsymbol{x}$. For large B, $vol(S_i) \vec{\approx} \frac{1}{2^B \lambda(\bar{\boldsymbol{x}}_i)}$, $P_i \vec{\approx} f(\bar{\boldsymbol{x}}_i) vol(S_i)$, and

$$E_d(Q_{opt}) \approx \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}} \frac{nf(\bar{\boldsymbol{x}}_i) vol(S_i)}{2(n+2)} \left(\frac{|\boldsymbol{D}(\bar{\boldsymbol{x}}_i)|}{(2^B \lambda(\bar{\boldsymbol{x}}_i) \kappa_n)^2} \right)^{1/n}$$
$$\vec{\approx} \frac{n 2^{-2B/n} \kappa_n^{-2/n}}{2(n+2)} \int_{\boldsymbol{x}} \lambda^{-2/n}(\boldsymbol{x}) |\boldsymbol{D}(\boldsymbol{x})|^{1/n} f(\boldsymbol{x}) d\boldsymbol{x}. \quad (7)$$

The function $\lambda(\boldsymbol{x})$ which minimizes the above approximation, subject to the unit integral constraint, can be found using standard Lagrange multiplier techniques, or using Holder's inequality [8,9,11]. By substituting the optimal $\lambda(\boldsymbol{x})$ into equation 7, the minimum error, achieved using the optimal quantizer, can be approximated as

$$E_d(Q_{opt}) \approx \frac{n 2^{-2B/n} \kappa_n^{-2/n}}{2(n+2)} \times \left(\int_{\boldsymbol{x}} (|\boldsymbol{D}(\boldsymbol{x})|^{1/n} f(\boldsymbol{x}))^{n/(n+2)} d\boldsymbol{x} \right)^{(n+2)/n}. \quad (8)$$

In summary, training a high rate vector quantizer by minimizing a distortion measure $d(\boldsymbol{x}, Q(\boldsymbol{x}))$, results in quantization points with an optimal density, Voronoi region shapes approximated by $T(\boldsymbol{D}(\boldsymbol{x}), (2^B \lambda_{opt}(\boldsymbol{x}))^{-1})$, and an expected distortion approximated (in lower bound) by equation 8. For squared Euclidean distance, $\boldsymbol{D}(\boldsymbol{x}) = 2\boldsymbol{I}$ and the above equations are equal to the expressions given in [6].

Suboptimal Distortion Measures

This section derives expressions for the performance of a high rate vector quantizer, as measured by distortion measure $d_2(\boldsymbol{x}, Q(\boldsymbol{x}))$, which has been trained by minimizing another distortion measure $d_1(\boldsymbol{x}, Q(\boldsymbol{x}))$. d_2 can be viewed as the distortion measure of relevance, and d_1 can be viewed as the distortion measure used in training and quantization due to ease of implementation. Here, $Q_1(\boldsymbol{x})$ is the quantization mapping resulting from training a vector quantizer by minimizing d_1.

The expression for the distortion incurred in the resulting quantization, as measured by d_2, using the previous steps, can be shown to be

$$E_{d_2}(Q_1) = \int_{\boldsymbol{x}} d_2(\boldsymbol{x}, Q_1(\boldsymbol{x})) f(\boldsymbol{x}) d\boldsymbol{x}$$
$$\vec{\approx} \frac{1}{2} \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}_1} f(\bar{\boldsymbol{x}}_i) \int_{\boldsymbol{x} + \bar{\boldsymbol{x}}_i \in S_i} \boldsymbol{x}^T \boldsymbol{D}_2(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x}. \quad (9)$$

where $\boldsymbol{D}_2(\bar{\boldsymbol{x}})$ is the sensitivity matrix with respect to the distortion measure d_2. The function $\boldsymbol{D}_1(\boldsymbol{x})$ is similarly defined for d_1. Using the fact that the approximate Voronoi region shape S_i defined by $Q_1(\boldsymbol{x})$ is given by $T(\boldsymbol{D}_1(\bar{\boldsymbol{x}}_i), vol(S_i))$,

$$\int_{\boldsymbol{x} + \bar{\boldsymbol{x}}_i \in S_i} \boldsymbol{x}^T \boldsymbol{D}_2(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x} \approx \int_{\boldsymbol{x} \in T(\boldsymbol{D}_1(\bar{\boldsymbol{x}}_i), vol(S_i))} \boldsymbol{x}^T \boldsymbol{D}_2(\bar{\boldsymbol{x}}_i) \boldsymbol{x} d\boldsymbol{x}$$
$$= \frac{vol(S_i)}{n+2} \left(\frac{vol^2(S_i) |\boldsymbol{D}_1(\bar{\boldsymbol{x}}_i)|}{\kappa_n^2} \right)^{1/n} tr\left(\boldsymbol{D}_1^{-1}(\bar{\boldsymbol{x}}_i) \boldsymbol{D}_2(\bar{\boldsymbol{x}}_i) \right).$$

The result shown in [8,9] was used to proceed to the final line in the above expression. Substituting this result into equation 9 and using the approximation for $f(\boldsymbol{x})$ used in equation 5 results in

$$E_{d_2}(Q_1) \approx \sum_{\bar{\boldsymbol{x}}_i \in \bar{\boldsymbol{X}}_1} \frac{P_i tr\left(\boldsymbol{D}_1^{-1}(\bar{\boldsymbol{x}}_i) \boldsymbol{D}_2(\bar{\boldsymbol{x}}_i) \right)}{2(n+2)} \left(\frac{vol^2(S_i) |\boldsymbol{D}_1(\bar{\boldsymbol{x}}_i)|}{\kappa_n^2} \right)^{1/n}.$$

Approximating the sum as an integral as before yields

$$E_{d_2}(Q_1) \approx \frac{2^{-2B/n} \kappa_n^{-2/n}}{2(n+2)} \times \int_{\boldsymbol{x}} \lambda_1^{-2/n}(\boldsymbol{x}) |\boldsymbol{D}_1(\boldsymbol{x})|^{1/n} tr\left(\boldsymbol{D}_1^{-1}(\boldsymbol{x}) \boldsymbol{D}_2(\boldsymbol{x}) \right) f(\boldsymbol{x}) d\boldsymbol{x},$$

where $\lambda_1(\boldsymbol{x})$ is the optimal codevector density function found by minimizing $d_1(\boldsymbol{x}, Q(\boldsymbol{x}))$ in training. Substituting in the expression for $\lambda_1(\boldsymbol{x})$ results in the following final approximation:

$$E_{d_2}(Q_1) \approx \frac{2^{-2B/n} \kappa_n^{-2/n}}{2(n+2)} \times \frac{\int_{\boldsymbol{x}} \left(|\boldsymbol{D}_1(\boldsymbol{x})|^{1/n} f(\boldsymbol{x}) \right)^{n/(n+2)} tr\left(\boldsymbol{D}_1^{-1}(\boldsymbol{x}) \boldsymbol{D}_2(\boldsymbol{x}) \right) d\boldsymbol{x}}{\left(\int_{\boldsymbol{x}} \left(|\boldsymbol{D}_1(\boldsymbol{x})|^{1/n} f(\boldsymbol{x}) \right)^{n/(n+2)} d\boldsymbol{x} \right)^{-2/n}}.$$

If $\boldsymbol{D}_1(\boldsymbol{x}) = \boldsymbol{D}_2(\boldsymbol{x})$, the above expression is equal to the result given in equation 8.

Practical Implications

At high rates, the results of the previous sections show that only the second order terms in the Taylor series expansion of the distortion measure are relevant. Thus, if a high rate vector quantizer is trained by minimizing the expected value of the quadratic distortion measure

$$\hat{d}(\boldsymbol{x}, Q(\boldsymbol{x})) = \frac{1}{2}(\boldsymbol{x} - Q(\boldsymbol{x}))^T \boldsymbol{D}(\boldsymbol{x})(\boldsymbol{x} - Q(\boldsymbol{x})), \quad (10)$$

the high rate codevector density, the Voronoi region shapes, and the performance will approach that achieved by a vector quantizer which has been trained by minimizing the "true" distortion measure, $d(\boldsymbol{x}, Q(\boldsymbol{x}))$. There are several advantages to be gained by using $\hat{d}$ rather than d in training and quantizing. First, it is easy to compute the centroid of a quadratic measure such as in equation 10, whereas it may be impossible to compute the centroid of the original error measure. This allows the quantizer to be built using the generalized Lloyd or LBG algorithms. Without the existence of an expression for the centroid of a region, constructing a quantizer becomes difficult or impossible. Second, the complexity incurred in computing the distortion measure for each codevector in $\bar{\boldsymbol{X}}$ during quantization may be much less for the quadratic distortion measure than for the true distortion measure. To quantize using the quadratic measure, the matrix $\boldsymbol{D}(\boldsymbol{x})$ must be computed only once per input vector.

For some applications, computing the full quadratic distortion for each vector in $\bar{\boldsymbol{X}}$ during quantization may still be too complex, and the use of simple MSE or WMSE distortion measures may be desirable. In [8,9], the theoretical expressions presented above were evaluated for simple MSE and WMSE measures, using the substitutions $\boldsymbol{D}_{MSE}(\boldsymbol{x}) = 2\boldsymbol{I}$ and $\boldsymbol{D}_{WMSE}(\boldsymbol{x}) = 2diag(w_1(\boldsymbol{x}), ..., w_n(\boldsymbol{x}))$ for $\boldsymbol{D}_1(\boldsymbol{x})$, respectively. Using Lagrange multiplier and calculus of variations techniques, it can be shown that the optimal weightings for a WMSE measure are the diagonal elements of the sensitivity matrix, i.e. $w_i(\boldsymbol{x}) = (\boldsymbol{D}_2(\boldsymbol{x}))_{i,i}$.

High Rate LPC Quantization

To apply the results from the previous sections to the problem of LPC quantization, the sensitivity matrices must be determined for the various different LPC parameter representations. In the previous sections, d_2 was the distortion measure of interest. For LPC quantization, d_2 is the LSD measure, and the sensitivity matrices with respect to the LSD measure must be computed.

Denote the set of filter taps corresponding to a vth order LPC filter by the vector

$$\boldsymbol{a} = [a_1 a_2 \ldots a_v]^T.$$

Most systems for compressing telephone bandwidth speech use a linear prediction filter of 10th order, i.e. $v = 10$. Note that the "0-th" tap, normally constrained to be equal to one in LPC literature, is not included in this vector. The LSD in dB^2 incurred by quantizing the vector $\boldsymbol{a}$ to a vector $\bar{\boldsymbol{a}}$ is given by

$$LSD(\boldsymbol{a}, \bar{\boldsymbol{a}}) = \frac{\beta}{2\pi}\int_{-\pi}^{\pi} \left(\ln(|A(\omega)|^2) - \ln(|\bar{A}(\omega)|^2)\right)^2 d\omega.$$

Here $\beta = (10/\ln(10))^2$, $A(\omega) = 1 - \sum_{i=1}^{v} a_i e^{j\omega i}$, and $\bar{A}(\omega) = 1 - \sum_{i=1}^{v} \bar{a}_i e^{j\omega i}$.

The computation and properties of the sensitivity matrices for LPC coefficients, reflection coefficients, LAR parameters, arcsine parameters, and LSP frequencies are discussed in other related work [8–10]. In these works, closed form expressions, involving no integration, are given for the sensitivity matrices of all of the above mentioned parameter sets, allowing for efficient computation of the optimal weightings in real-time on current DSP chips.

Some interesting results from [8–10] are:

- For the LPC coefficients, the sensitivity matrix is proportional to the autocorrelation matrix $\boldsymbol{R}_A$, familiar in linear prediction theory. Hence, at high rate, the LSD measure approaches the standard LPC prediction error measure.
- For reflection coefficients, LAR parameters, and arcsine parameters, the sensitivity matrix has smaller relative off-diagonal elements than $\boldsymbol{R}_A$, showing that diagonal approximations for the sensitivity matrix in the reflection coefficient domain are better than in the LPC domain. However, off-diagonal components still exist. For arcsine and LAR parameters, there is a smaller spread in the diagonal elements of the sensitivity matrix (i.e. the scalar sensitivities of the parameters), compared to those of the reflection coefficients.
- For the LSP frequencies, *the sensitivity matrix is exactly diagonal*, meaning that at high rate, the LSD measure is equivalent to a WMSE measure in the LSP frequencies! This is true only of the LSP frequencies and thus gives theoretical justification for using WMSE measures of LSP frequencies for VQ of LPC parameters. This implies that training a high rate VQ using the sensitivity weighted WMSE-LSP measure will produce optimal results.

Theory vs. Experiment

In this section, experimental results are given and compared with the theoretical results from the previous sections. The LPC coefficients were determined from 8kHz sampled speech using the procedure in the TIA IS-96 standard [13]. A database of over 160,000 frames of speech was used for VQ training. This database was taken from a larger database of speech, and the silence between the spoken speech was removed using an adaptive algorithm. Approximately half of the database was recorded with a flat spectral recording response, and the other half of the database was IRS filtered. For testing, a separate 20,000 frames of speech was used.

To compare the theoretical predictions with experimental results, a split-VQ (or partitioned-VQ) scheme was investigated, for pragmatic reasons explained in [8, 9]. The split-VQ scheme tested here divided the LSP vectors into 3 subvectors, with the first subvector containing the 3 lowest LSPs, the second subvector containing the 3 middle LSPs, and the final subvector containing the 4 highest LSPs. Separate vector quantizers were then constructed for each subvector. For the overall VQ, an equal number of bits was allocated for each subvector.

A theoretical approximation to the overall high rate VQ performance can be found by summing the theoretical approximations for the 3 subvector quantizers, since the LSD is a weighted sum of the errors in the individual LSP parameters at high rate. To evaluate the theoretical approximations, the density functions of the subvectors must be estimated. For each subvector, a multi-dimensional histogram approach with non-uniform bin sizes was computed from the training set [8, 9]. The non-uniform bin size was used to prevent a "slope overload" effect at the histogram bins corresponding to the tails of the distribution, while maintaining sufficient resolution in the histogram bins corresponding to LSP subvectors where the sensitivity matrix function changes most rapidly (i.e. where LSP frequencies are close to each other). Because the sensitivity matrix for each subvector is not a function solely of the LSP frequencies in the subvector, the average of the sensitivity matrices in each histogram bin was used to estimate the expected sensitivity matrix in that bin. Numerical integration over the density functions of the equations in the previous sections was then used to obtain approximate theoretical performances.

Vector Quantizer Results

Figure 1 shows the performance of the overall vector quantizers. The solid line shows the performance of the VQ trained by minimizing the sensitivity weighted WMSE measure. The dashed line shows the performance of the VQ trained by minimizing the MSE measure. The dotted line shows the theoretical approximate performance of the WMSE vector quantizer, computed as described above. The dash-dot line shows the theoretical approximate performance of the MSE vector quantizer. As can be seen, the theoretical approximations are fairly accurate for distortions under 1-2 dB^2, but are slightly lower than the experimental performances at high rate due to the lower bound used in the approximations. The use of the simpler MSE measure requires a quantizer which has approximately 2.5 extra bits to achieve the same performance as a quantizer using the optimal WMSE measure. The performance of the optimal sensitivity WMSE measure was also compared with the performance using the previously proposed WMSE measures used in [2,3], and was shown to be about 1 bit better than the performance achieved using these suboptimal weightings [8,9].

The overall performance of the vector quantizers reported here is somewhat worse than that reported in other previous work for several reasons. First, the databases were compiled from two different sets of source material, creating a broader spread of the LSP density function and thus a harder source to encode. Second, the databases used here consisted of active speech only and the LPC vectors corresponding to silence/background noise, which are relatively insensitive to quantization error, were removed. Third, the full LSD, averaged over the entire frequency range was

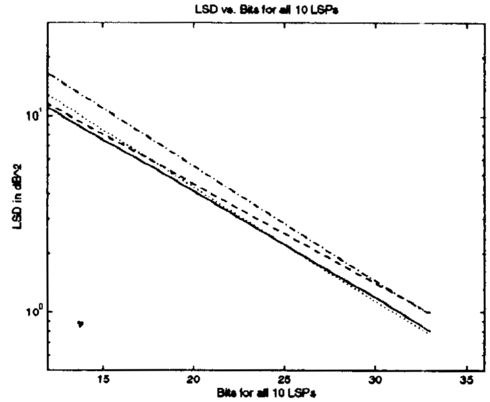

Figure 1: Theoretical and Experimental Performance of MSE and WMSE based LSP VQs

used, as opposed to the LSD averaged over only the lower 3/4 of the frequency range used in some previous work (e.g. [2]). Finally, the LSD was computed as the average the dB^2 values for each frame, rather than as the average of the dB values.

References

[1] A. Gersho and R. Gray, *Vector Quantization and Signal Compression.* Kluwer Academic Press, 1990.

[2] K. K. Paliwal and B. Atal, "Efficient Vector Quantization of LPC Parameters at 24 bits/frame," *IEEE Trans. on SAP*, vol. 1, Jan. 1993.

[3] R. Laroia, N. Phamdo, and N. Farvardin, "Robust and Efficient Quantization of Speech LSP Parameters using Structured Vector Quantizers," in *Proceedings of ICASSP*, 1991.

[4] F. K. Soong and B.-H. Juang, "Optimal Quantizaton of LSP Parameters," in *Proceedings of ICASSP*, 1988.

[5] B. Atal, R. Cox, and P. Kroon, "Spectral Quantization and Interpolation for CELP Coders," in *Proceedings of ICASSP*, 1989.

[6] A. Gersho, "Asymptotically Optimum Block Quantization," *IEEE Trans on IT*, vol. 25, July 1979.

[7] Y. Yamada, S. Tazaki, and R. Gray, "Asymptotic Performance of Block Quantizers with Difference Distortion Measures," *IEEE Trans. on IT*, vol. 26, Jan. 1980.

[8] W. R. Gardner and B. D. Rao, "Theoretical Analysis of the High Rate Vector Quantization of LPC Parameters," *submitted to IEEE Trans. on Speech and Audio*, 1994.

[9] W. R. Gardner, *Modeling and Quantization Techniques for Speech Compression Systems.* PhD thesis, University of California, San Diego, 1994.

[10] W. R. Gardner and B. D. Rao, "Optimal Distortion Measures for the High Rate Vector Quantization of LPC Parameters," *submitted to ICASSP*, 1995.

[11] R. M. Gray, *Source Coding Theory.* Kluwer Academic Publishers, 1990.

[12] J. Conway and N. Sloane, "Voronoi Regions of Lattices, Second Moments of Polytopes, and Quantization," *IEEE Trans. on IT*, vol. 28, Mar. 1982.

[13] "Speech Service Option Standard for Wideband Spread Spectrum Digital Cellular System - EIA/TIA/IS-96," 1994.