Incorporating Error Shaping Technique into LSF Vector Quantization

Hsi-Wen Nein and Chin-Teng Lin, *Senior Member, IEEE*

***Abstract*—This paper presents an error shaping technique for line spectrum frequency (LSF) vector quantization. The error shaping technique based on the weighted logarithm spectral distortion (WLSD) measure can be used for shaping the spectral distortion distribution of quantization error into any different curve depending on what kind of weighting function is used. However, the high computational complexity of the WLSD measure deters this error shaping technique from practical use. To solve this problem, we approximate the WLSD measure by the quadratically weighted measure or the weighted mean squared error (WMSE) measure and propose an optimal error shaping technique of LSF vector quantization. In this proposed error shaping technique, the optimal WMSE weights (i.e., the optimal weights of LSF parameters) are determined based on the theoretical analysis of the WLSD measure. Three experiments are performed to check the performance of the proposed error shaping technique. One experiment is set up by incorporating human perception into the LSF quantization and another is set up by emphasizing the human-sensitivity frequency band in lower frequency bandwidth 0–3 kHz. In the third experiment, we apply the proposed error shaping technique to the LSF quantization of a CELP coder to test how it affects the overall speech quality in an actual speech coding algorithm.**

***Index Terms*—Human perception, quadratically weighted measure, sensitivity matrix, weighted logarithm spectral distortion (WLSD), weighted mean squared error (WMSE).**

I. Introduction

THE LINEAR predictive coding (LPC) model for speech signals is used in many modern speech compression systems. In these systems, the LPC filter coefficients are usually transformed to line spectrum frequency (LSF) parameters which give a very effective representation for the quantization of LPC information [1], [2]. An example is the new 2400 bps MELP coder which uses 25 bit, multistage 10 LSF parameter quantization [3].

Most vector quantization (VQ) research focuses on minimizing the number of bits needed to represent the speech spectrum for achieving a low log-spectral distortion (LSD) between the unquantized speech spectrum and the quantized one. Although LSD is a common spectral distortion measure for evaluating the performance of the quantizers, it is difficult to design a quantizer that directly minimizes the overall LSD due to the high computational complexity of the LSD measure. Therefore, simpler distortion measures are designed for practical applications, such as the mean squared error (MSE) measure or weighted mean squared error (WMSE) measure for the LSF parameters.

Most LSF vector quantizers based on WMSE measure experimentally determine the weights of LSF parameters to minimize LSD [5]–[7]. The experimentally determined weights of LSF parameters used in these quantizers lack any theoretical analysis. Recently, Gardner and Rao [4] have proved that a WMSE measure of LSF parameters converges to the LSD measure in high-bit-rate VQ systems. In [4], the optimal WMSE weights of LSF parameters are determined based on the theoretical analysis of the LSD measure. However, one obvious drawback of the LSD measure is that it does not utilize the frequency-dependent property; i.e, all frequencies have equal weights in the LSD measure.

Since all frequencies are weighted equally in the LSD measure, the optimal weights of LSF parameters obtained in [4] does not take into account the perceptual property of human ear. To make better use of this property, Paliwal and Atal achieved good VQ performance using the WMSE measure with weights determined experimentally [5]. Cohn and Collua determined weights experimentally such that a spectral distribution of quantization error was perceived to be "balanced," i.e., the error at all frequencies contribute equally on average to the perceived distortion [7]. In summary, the weights determined in [5], [7] experimentally lack any theoretical analysis, and the optimal weights obtained from the theoretical analysis of the LSD measure [4] do not use the human perceptual characteristics.

In order to make better use of the perceptual property of human ear, we need use the weighted LSD (WLSD) measure instead of the unweighted LSD measure to calculate the distortion between the unquantized speech spectrum and the quantized one. In this way, the error distortion distribution of a quantizer can be shaped into any curve depending on which weighting function is used in the WLSD measure. If the chosen weighting function matches human perceptual characteristics, then the error distortion distribution of the quantizer and the perceived distortion will match in each frequency. Nevertheless, the simple measures such as MSE and WMSE are still required, since it is quite difficult, if not impossible, to design a quantizer that directly minimizes the overall WLSD due to the high computational complexity of the WLSD measure.

In this paper, we study the error shaping technique of LSF VQ based on the WLSD measure for any given weighting function by extending Gardner's work [4] to include the more general case of weighting LSD. Our theoretical analysis shows that

Manuscript received February 19, 1999; revised April 14, 2000. This research was supported by Lee and MTI Center for Networking Research and the National Science Council, R.O.C., under Grant NSC 89-2213-E-009-114. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Peter Kroon.

The authors are with the Department of Electrical and Control Engineering National Chiao-Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: ctlin@fnn.cn.nctu.edu.tw).

Publisher Item Identifier S 1063-6676(01)00844-6.

the quantization distortion with the WLSD measure is close to that of a quadratically weighted distortion measure, which is much lower in computational complexity. We also obtain the optimal weights for LSF parameter quantization from the theoretical analysis of the WLSD measure, since simple measure such as WMSE may be desirable for practical applications. In our experiments, we will show that quantizers using theoretically derived weights outperform quantizers using heuristically derived weights.

This paper is organized as follows. Section II introduces the error shaping technique using the WLSD measure, and then derives approximate solutions to the WLSD measure for the LSF VQ schemes that use a quadratic distortion measure or a WMSE distortion measure for training and quantizing. In Section III, two experimental LSF VQ results, and experimental testing results of applying the proposed technique to the LSF quantization of a CELP coder are presented. Finally, conclusions are drawn in Section IV.

II. Theoretical Analysis of Error Shaping

In this section, we first introduce an error shaping technique based on the WLSD measure. Then, the optimal weights used in the quadratic and WMSE distortion measures are derived in Section II-B. An algorithm of computing the weights used the quadratic and WMSE distortion measures is given in Section II-C.

A. Error Shaping Technique

The LSD measure, D, being a common measure of LPC performance is defined as follows [8]:

$$D^2(\mathbf{a}, \hat{\mathbf{a}}) = \frac{1}{\pi}\int_0^{\pi} [10\log_{10}(P(w)) - 10\log_{10}(\hat{P}(w))]^2\, dw \tag{1}$$

where w is the radian frequency, and $P(w)$ and $\hat{P}(w)$ are the linear prediction power spectra before and after quantization, respectively,

$$P(w) = \left|1 - \sum_{k=1}^{v} a_k e^{-jwk}\right|^{-2}$$
$$\hat{P}(w) = \left|1 - \sum_{k=1}^{v} \hat{a}_k e^{-jwk}\right|^{-2} \tag{2}$$

where v is the order of LPC filter and is constrained to be even, $\mathbf{a} = [a_1, a_2, \cdots, a_v]^T$ and $\hat{\mathbf{a}} = [\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_v]^T$ are the linear prediction coefficients corresponding to the unquantized and quantized LSF parameters, respectively. This measure does not account for human perceptual characteristics, since it gives equal weight to the error regardless of where it occurs in frequency.

To consider frequency-dependent error, the weighted LSD (WLSD) measure is defined as follows:

$$D_B^2(\mathbf{a}, \hat{\mathbf{a}}) = \frac{1}{2\pi B_0}\int_{-\pi}^{\pi} B^2(w) \cdot [10\log_{10}(P(w)) - 10\log_{10}(\hat{P}(w))]^2\, dw \tag{3}$$

where B_0 is a normalization factor making $B(w)/B_0$ to be unity root mean square (RMS). The values of $B(w)$ are all real and positive. The function $B(w)$ is called weighting function, which is required to be symmetric, i.e., $B(-w) = B(w)$. This constraint for the weighting function is reasonable since the interested frequency range is always in $[0, \pi]$.

The frequency-dependent weighting function $B(w)$ can be used for determining frequency-dependent error shape in frequency domain and thus, it can be assigned freely for different purposes. The error shaping technique with the WLSD measure for a given weighting function can be applied to any LSF quantizer which uses frequency-dependent error distortion measure. The WLSD measure is generalized from [11] in which the perceptual-based distortion measure is used.

B. Theoretical Analysis

Although the WLSD measure makes better use of the perceptual property of human ear, it is difficult to use in a LSF quantizer directly due to the high computational complexity of the WLSD measure. Hence, simpler distortion measures that approximate the WLSD measure are desired, such as the quadratic distortion measure and WMSE measure.

In Sections II-B1 and II-B2, we shall derive the approximation solutions of the WLSD measure for the LSF VQ schemes that use a quadratic distortion measure and a WMSE distortion measure for training and quantization, respectively. After deriving the approximation solutions of the WLSD measure, the theoretical analysis on high-rate VQ using WLSD measure is verified in Section II-B3.

1) Replacement of the WLSD Measure by the Quadratic Distortion Measure: In this section, we shall study the replacement of the WLSD measure by a quadratic distortion measure for computational simplicity. We first introduce a theorem in [4] showing that any continuously differentiable distortion function which satisfies some properties can be approximated by a quadratic distortion function to any desired degree for short distances.

Theorem 1: Let $\mathbf{x}$ be an n-dimensional vector and assume that a vector quantizer maps $\mathbf{x}$ to an output vector $\hat{\mathbf{x}}$, and $d(\mathbf{x}, \hat{\mathbf{x}})$ is a continuously differentiable distortion function having the following properties:

- $d(\mathbf{x}, \hat{\mathbf{x}}) \geq 0$ with equality holding only if $\mathbf{x} = \hat{\mathbf{x}}$;
- $d(\mathbf{x}, \hat{\mathbf{x}})$ having continuous partial derivatives of third order almost everywhere;
- second-order derivatives of $d(\mathbf{x}, \hat{\mathbf{x}})$, $(\partial^2 d(\mathbf{x}, \hat{\mathbf{x}}))/(\partial x_j \partial x_k)$, for all j, k, being positive definite almost everywhere.

Then, this $d(\mathbf{x}, \hat{\mathbf{x}})$ measure can be approximated exactly by a quadratic distortion measure, i.e.,

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}})^T \mathbf{D}(\mathbf{x})(\mathbf{x}-\hat{\mathbf{x}}) \qquad \text{for small distances} \tag{4}$$

where $\mathbf{D}(\mathbf{x})$ is an n by n dimensional matrix with its (j, k)th element defined by

$$D_{j,k}(\mathbf{x}) = \left.\frac{\partial^2 d(\mathbf{x}, \hat{\mathbf{x}})}{\partial x_j \partial x_k}\right|_{\hat{\mathbf{x}}=\mathbf{x}} \tag{5}$$

and is denoted as the "sensitivity matrix," since its elements represent the relative sensitivity of quantizing the various parameters.

Proof: See [9].

In general, the sensitivity matrix used in the quadratic distortion measure in (4) is not diagonal. The quadratic distortion

measure provides a simple method to computing the centroid of a quadratic measure, whereas it may be impossible to efficiently compute the centroid of the "true" distortion measure (such as the LSD or WLSD measure). Thus this allows the quantizer to be built using the generalized Lloyd or LBG algorithms [12].

It is readily apparent that Theorem 1 holds for the WLSD measure in (3) only if the weighting function $B(w)$ has continuous third-order derivatives, since the WLSD measure satisfies the properties required by $d(\mathbf{x}, \hat{\mathbf{x}})$ in Theorem 1. Hence, we can use the quadratic distortion measure in the right hand side of (4) to replace the WLSD measure in designing an LSF quantizer. To design an LSF quantizer based on the quadratic distortion measure, we need to determine the sensitivity matrix in (4). This issue is what we focus on in the rest of this subsection.

For analysis convenience, we rewrite (3) as follows:

$$\begin{aligned}&\mathrm{WLSD}(\mathbf{a}, \hat{\mathbf{a}})\\ &\quad = \frac{\alpha}{2\pi}\int_{-\pi}^{\pi} B^2(w)[\ln(|A(w)|^2) - \ln(|\hat{A}(w)|^2)]^2\, dw,\end{aligned} \tag{6}$$

where $\alpha = (10/\ln(10))^2/B_0$, $A(w) = 1 - \sum_{k=1}^{v} a_k e^{-jwk}$, $\hat{A}(w) = 1 - \sum_{k=1}^{v} \hat{a}_k e^{-jwk}$.

To derive the sensitivity matrix for LSF parameters based on (6), we first find out the sensitivity matrix for LPC parameters based on (6) and then apply Jacobian matrix of transformation which transforms LSF parameters to LPC parameters to it to obtain the sensitivity matrix for LSF parameters.

a) Sensitivity Matrix for LPC Parameters: The following theorem gives the exact form of the sensitivity matrix for LPC parameters.

Theorem 2: Let $h[n]$ and $b_s[n]$ denote the impulse responses of the discrete-time filters $1/A(w)$ and $B_s(w) \triangleq B^2(w)$ in (6), respectively. Notice that $h[n]$ is a causal signal and $b_s[n]$ is a symmetric and noncausal signal. Then, the elements of the sensitivity matrix for LPC parameters are

$$\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{a}, \hat{\mathbf{a}})}{\partial \hat{a}_k \partial \hat{a}_l}\right|_{\hat{\mathbf{a}}=\mathbf{a}} = 4\alpha R_{Am}(k, l) \tag{7}$$

where

$$R_{Am}(k, l) = \sum_{n=-\infty}^{\infty} b_s[n](r_c[(k+l)-n] + r[(k-l)-n]) \tag{8}$$

where $r[n]$ is the traditional autocorrelation function of the impulse response $h[n]$, i.e.,

$$r[n] = \sum_{m=0}^{\infty} h[m]h[m+n] \tag{9}$$

and $r_c[n]$ is the convolution of $h[-n]$ with $h[-n]$, i.e.,

$$\begin{aligned} r_c[n] &= \sum_{m=0}^{\infty} h[-m]h[-(n-m)] = \sum_{m=0}^{\infty} h[-m]h[m-n]\\ &= \sum_{m=n}^{0} h[-m]h[m-n], \qquad n \le 0. \end{aligned} \tag{10}$$

Proof: See Appendix A.

From the results in Theorem 2, the sensitivity matrix for LPC parameters can be written as

$$\mathbf{D}_A(\mathbf{a}) = 4\alpha \mathbf{R}_{Am} \tag{11}$$

where $\mathbf{R}_{Am}$ is defined as

$$\mathbf{R}_{Am} = \begin{bmatrix} R_{Am}(1, 1) & \cdots & R_{Am}(1, v)\\ \vdots & \ddots & \vdots\\ R_{Am}(v, 1) & \cdots & R_{Am}(v, v)\end{bmatrix}. \tag{12}$$

The matrix $\mathbf{D}_A(\mathbf{a})$ has two properties listed in the following:

a) $\mathbf{D}_A(\mathbf{a})$ is a symmetric matrix since $r[n]$ and $b_s[n]$ are symmetric signals. The proof is given by the following equalities:

$$\begin{aligned} R_{Am}(i, j) &= \sum_{n=-\infty}^{\infty} b_s[n](r_c[i+j-n] + r[(i-j)-n])\\ &= \sum_{n=-\infty}^{\infty} b_s[n](r_c[j+i-n] + r[-(i-j)+n])\\ &= \sum_{n=-\infty}^{\infty} b_s[n]r_c[j+i-n] + \sum_{n=-\infty}^{\infty} b_s[-n]r[(j-i)-n]\\ &= R_{Am}(j, i). \end{aligned} \tag{13}$$

b) $\mathbf{D}_A(\mathbf{a})$ is positive-definite almost everywhere from Theorem 1.

Property a) is useful for saving the computation time in the calculation of $\mathbf{D}_A(\mathbf{a})$.

b) Sensitivity Matrix for LSF Parameters: In order to define LSF parameters, the inverse filter, $A(z)$, is used to construct two polynomials:

$$P(z) = A(z) + z^{-(v+1)}A(z^{-1}) \tag{14}$$

and

$$Q(z) = A(z) - z^{-(v+1)}A(z^{-1}). \tag{15}$$

The roots of these polynomials are usually called LSF parameters. Some important properties of LSF parameters are detailed in [2], [5].

Denote the transformation for LPC vector $\mathbf{a}$ to the LSF parameters set $\mathbf{w} = [w_1, w_2, \cdots, w_v]^T$ by the function $\mathbf{w}(\mathbf{a})$ and the reverse transformation by $\mathbf{a}(\mathbf{w})$. Then, the element of the sensitivity matrix $\mathbf{D}_w(\mathbf{w})$ for the LSF parameters is given by

$$\begin{aligned}&\left.\frac{\partial \mathrm{WLSD}(\mathbf{a}(\mathbf{w}), \mathbf{a}(\hat{\mathbf{w}}))}{\partial \hat{w}_k \partial \hat{w}_l}\right|_{\hat{\mathbf{w}}=\mathbf{w}}\\ &\quad = \sum_{m=1}^{v}\sum_{n=1}^{v} \frac{\partial a_m(\hat{\mathbf{w}})}{\partial \hat{w}_k}\frac{\partial a_n(\hat{\mathbf{w}})}{\partial \hat{w}_l}\\ &\qquad \cdot \left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{a}(\mathbf{w}), \hat{\mathbf{a}})}{\partial \hat{a}_m \partial \hat{a}_n}\right|_{\hat{\mathbf{w}}=\mathbf{w},\, \hat{\mathbf{a}}=\mathbf{a}(\mathbf{w})}\\ &\quad = \left(\mathbf{J}_w^T(\mathbf{w})\mathbf{D}_A(\mathbf{a})\mathbf{J}_w(\mathbf{w})\right)_{k,l}\end{aligned} \tag{16}$$

where $\mathbf{D}_A(\mathbf{a})$ is given in (11), and $\mathbf{J}_w(\mathbf{w})$ is the Jacobian matrix of the transform $\mathbf{a}(\mathbf{w})$, which has its (j, k)th element defined by

$$(\mathbf{J}_w(\mathbf{w}))_{j,k} = \frac{\partial a_j(\mathbf{w})}{\partial w_k}. \tag{17}$$

By substituting the result in (11) into (16), we can rewrite the sensitivity matrix of LSF parameters in (16) as

$$\mathbf{D}_w(\mathbf{w}) = \mathbf{J}_w^T(\mathbf{w})\mathbf{R}_{Am}(\mathbf{a})\mathbf{J}_w(\mathbf{w}). \tag{18}$$

2) Approximation of the WLSD Measure by the WMSE Measure: In some applications, computing the full quadratic distortion measure for each vector in a codebook during quantization may still be too complex, and the use of relatively simple WMSE distortion measure may be desirable. The WMSE distortion measure is a special case of quadratic distortion measure with a diagonal matrix, which is defined as

$$\begin{aligned} d_{\mathrm{WMSE}}(\mathbf{x}, \hat{\mathbf{x}}) &= (\mathbf{x}-\hat{\mathbf{x}})^T \begin{bmatrix} \omega_1(\mathbf{x}) & & & \mathbf{0} \\ & \omega_2(\mathbf{x}) & & \\ & & \ddots & \\ \mathbf{0} & & & \omega_v(\mathbf{x}) \end{bmatrix} (\mathbf{x}-\hat{\mathbf{x}}) \\ &= \sum_{i=1}^{v} \omega_i(\mathbf{x})(x_i - \hat{x}_i)^2 \end{aligned} \tag{19}$$

where $\mathbf{x}$ and $\hat{\mathbf{x}}$ are column vectors of the original and quantized LSF parameters, and $\omega_i(\mathbf{x})$ is the ith weight corresponding to the ith element of the original vector, x_i. It is obvious that the quadratic distortion measure in (4) is equivalent to the WMSE measure when the sensitivity matrix in (4) is diagonal.

An important result showing that at high-bit rate, a WMSE measure in LSF quantization is equivalent to the LSD measure [4]. However, this does not hold when the WLSD measure instead of the LSD measure is concerned. The following theorem shows this point.

Theorem 3: The optimal sensitivity matrix for LSF parameters is nondiagonal if the length of impulse response of the weighting function $b_s[n]$ used in the WLSD measure is larger than three.

Proof: See Appendix B.

From Theorem 3, we understand that the quadratically weighted distortion measure in a LSF quantizer cannot reduce to the WMSE measure, since the sensitivity matrix in the former is not a diagonal matrix as required in the latter. Hence, from Theorem 3, there exists no WMSE measure that can be equivalent to the WLSD measure. However, the following theorem from [4] helps to find the optimal WMSE measure that is closest to the WLSD measure.

Theorem 4: The optimal weights in the WMSE measure are the diagonal elements of the sensitivity matrix using a continuously differentiable distortion measure having the same properties in Theorem 1.

Proof: See [4].

Since the WLSD measure is a continuously differentiable distortion measure and satisfies the properties in Theorem 1, it has the property shown in Theorem 4. In other words, the weights that make the WMSE measure approximate the WLSD measure optimally are the diagonal elements of the sensitivity matrix for LSF parameters, $\mathbf{D}_w(\mathbf{w})$ in (18). Notice that if the quadratic or WMSE distortion measure D_m is used for designing a LSF quantizer, then the factor α in the sensitivity matrix [see (11)] can be neglected, i.e., using $\overline{D}_m = \alpha D_m$ in training and quantization would produce identical results.

3) Verification of the Theoretical Analysis on High-Rate VQ Using WLSD Measure: The theoretical approximation, E_Q, for the performance of *high-rate* VQ schemes that use quadratic distortion measures for training and quantizing as described in Theorem 1 has been derived in [4]. In addition, based on LSD measure, an experiment was set up in [4] for comparing the result of the theoretical approximation to that of high-rate VQ trained and quantized using quadratic distortion measure. The training database of 160 000 frames of speech and the database of about 20 000 frames of speech were used for training and testing, respectively. The 10-dimensional vectors of LSF parameters and the LSF sensitivity matrix with respect to the LSD measure for each frame in the training and testing databases were computed. The reason for using LSF parameters as the input vector of the vector quantizer is that the LSF sensitivity matrix with respect to LSD measure is diagonal, which results in simpler computation. In other words, the quadratic distortion measure is equivalent to the WMSE distortion measure for a vector quantizer using the LSF sensitivity matrix with respect to the LSD measure to quantize LSF parameters. In [4], the theoretical performance of the high-rate VQ was computed by first estimating the density function of the set of LSF vectors using the 160 000 training frames, and then calculating its theoretical approximation E_Q with the density function in a numerical method. The experimental results in [4] have shown that the theoretical approximation (E_Q) approaches to the performance (LSD) of the high-rate VQ trained and quantized using the diagonal LSF sensitivity matrix.

Since the measure in our paper is based on the WLSD measure rather than the LSD measure, the comparisons between the theoretical approximation (E_Q) and the experimental result of the high-rate VQ using the (*nondiagonal*) LSF sensitivity matrix with respect to the WLSD measure obtained in Section II-B2 should be made for further verifying our work. The comparisons can be made by the same experiment set up described in [4] with the same training and testing databases except that the LSD measure is replaced by the WLSD measure. Thus, similar results to those presented in [4] can be obtained since the WLSD measure can be viewed as the warped LSD measure according to the following manipulation. By defining

$$C^2(w) = \frac{B^2(w)}{B_0} \tag{20}$$

we can obtain

$$\begin{aligned} \mathrm{WLSD} &= D_B^2 \\ &= \frac{1}{2\pi B_0} \int_{-\pi}^{\pi} B^2(w)[10\log_{10}(P(w)) \\ &\qquad - 10\log_{10}(\hat{P}(w))]^2\, dw \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}[10\log_{10}(P^{C(w)}(w)) \\
&\qquad - 10\log_{10}(\hat{P}^{C(w)}(w))]^2\,dw \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}[10\log_{10}(E(w)) - 10\log_{10}(\hat{E}(w))]^2\,dw \\
&= \text{warped LSD} \qquad (21)
\end{aligned}
$$

where $E(w) = P^{C(w)}(w)$ and $\hat{E}(w) = \hat{P}^{C(w)}(w)$ are the warped power spectra. From (21), we find that the power spectrum $P(w)$ in WLSD can be mapped into the warped power spectrum $E(w)$ in the warped LSD. After transforming the WLSD measure into the warped LSD measure in (21), we can obtain the LSF sensitivity matrix with respect to the warped LSD measure by the method described in [4]. According to the proof in [4], such a LSF sensitivity matrix is diagonal. Since the warped LSD measure in (21) and the original LSD measure used in [4] are in the exactly same form, and both of their respective LSF sensitivity matrices are diagonal, the result of experimental justification for the theoretical analysis on the performance of high-rate VQ in [4] can be applied to the WLSD case directly. We can thus verify that the theoretical approximation, E_Q, is very close to the experimental result of the high-rate VQ using the LSF sensitivity matrix with respect to the WLSD measure obtained in Section II-B2. In other words, the quadratic distortion measure of LSF parameters converges to the WLSD measure in a high-rate VQ system. Furthermore, according to Theorem 4, when the WMSE measure of LSF parameters is used, the WMSE measure using the diagonal elements of the LSF sensitivity matrix with respect to the WLSD measure in Section II-B2 is close to the WLSD measure optimally in the high-rate VQ system.

C. Computational Implementation

In order to compute the sensitivity matrix for LSF parameters given by (18), we need to determine firstly the sensitivity matrix for LPC parameters, $\mathbf{R}_{Am}$, and the Jacobian matrix of the transformation from LSF parameters to LPC parameters, $\mathbf{J}_w(\mathbf{w})$. The efficient computation algorithm of $\mathbf{J}_w(\mathbf{w})$ can be found in [4]. Thus, what remains is the computation of $\mathbf{R}_{Am}$.

Two problems exist in computing $\mathbf{R}_{Am}$; impulse response $h[n]$ of the LPC filter $1/A(z)$ is not a finite sequence, and $b_s[n]$ may be an infinity sequence. The first problem results in the difficulty that the autocorrelation function of $h[n]$ cannot be computed, since $h[n]$ is an infinity sequence. This problem can be solved by using the relation between the autocorrelation sequence ($R_A[k]$) of the windowed frame of input speech and the autocorrelation sequence ($R_s[k]$) of $h[n]$ as derived in [14]. However, in most cases, in order to avoid sharp spectral peaks in LPC spectrum which may result in unnatural synthesized speech, a fixed bandwidth expansion is applied to each pole of LPC coefficients, by replacing a_i by $a_i\gamma^i$ for $1 \le i \le v$ and $0 < \gamma < 1$. This results in the fact that the autocorrelation sequence ($R_r[k]$) of the impulse response ($h_r[n]$) of the bandwidth-expanded LPC ($A(z/\gamma)$) cannot be directly obtained from the autocorrelation sequence ($R_A[k]$) of the speech signal. Hence, for simply computing the autocorrelation sequence $R_r[k]$ from $h_r[n]$, we assume that $h_r[n]$ is zero after some large time index L, since $A(z/\gamma)$ is an autoregressive model. Under this assumption, the autocorrelation sequence $R_r[k]$ can be computed as

$$R_r[k] = \sum_{n=0}^{L-k} h_r[n]h_r[n+k]. \qquad (22)$$

This assumption is not very accurate and may cause some error in computing the autocorrelation sequence $R_r[k]$. Of course, if the fixed bandwidth expansion is not used, $R_r[k]$ can be directly and accurately obtained from the speech signal.

The second problem caused by infinity impulse response of $b_s[n]$ can also be fixed by selecting finite impulse response sequence to avoid the convolution of infinite sequence $b_s[n]$ and $r_c[n]$ in (8). If we only have the weighting function $B_s(w)$ in the frequency domain, then $b_s(n)$ can be obtained by taking the inverse discrete Fourier transform (IDFT) of $B_s(w)$. If $b_s[n]$ is truncated to $\overline{b}_s[n]$ for selecting finite impulse response with frequency response $\overline{B}_s(w)$, then $\overline{b}_s[n]$ must be a positive definite signal to guarantee that the resulting quadratic weighting matrix is positive definite. Here, we provide two simple methods to obtain positive-definite $\overline{b}_s[n]$ by using the common windows $w_t[n]$ in the design of FIR filters such as the rectangular window, Hanning window and Hamming window, etc., except that we put the center of the window at time index 0. One method to obtain the positive-definite $\overline{b}_s[n]$ is by applying a positive definite window, generated from the autocorrelation of $w_t[n]$, to $b_s[n]$. If the truncated $\overline{b}_s[n]$ is designed by this method, the weighting function used in the WLSD measure becomes

$$\overline{B}_s(w) = B_s(w) \otimes W_t^2(w),$$

where $W_t(w)$ denotes the frequency response of $w_t[n]$. The other method is firstly to compute the impulse response of $B(z) = B_s(z)/B(z^{-1})$, denoted as $b[n]$, by the spectral factorization procedure, where we assume $B_s(w)$ can be spectrally factorized. Then, by applying the window $w_t[n]$ to $b[n]$ for computing the finite impulse response $\overline{b}[n]$, the positive-definite $\overline{b}_s[n]$ can be obtained by calculating the autocorrelation of $\overline{b}[n]$. In the second method, the weighting function used in the WLSD measure is given by

$$\overline{B}_s(w) = |B(w) \otimes W_t(w)|^2.$$

The truncated signal $\overline{b}_s[n]$ obtained by these two methods is a positive-definite signal, since its Fourier transform $\overline{B}_s(w)$ is positive [10]. Hence, the quadratic weighting matrix obtained from Theorem 1 by applying $\overline{B}_s(w)$ to the WLSD measure is positive-definite almost everywhere.

With the above two problems being fixed, the sensitivity matrix for LPC parameters, $\mathbf{R}_{Am}$, can be computed as follows. Since the bandwidth expansion will be used in our experiments, we assume that $h[n]$ used for computing $r[n]$ in (9) is zero after a large time index L. In addition, assuming that $b_s[n]$ is zero outside the range $-N \le n \le N$, we can then rewrite Eqs. (9) and (10) as

$$r[k] = r[-k] = \sum_{n=0}^{L-k} h[n]h[n+k], \quad k = 0, 1, 2, \cdots, N+v-1 \qquad (23)$$

and

$$r_c[k] = \sum_{n=k}^{0} h[-n]h[n-k], \quad k = 0, -1, \cdots, -N+1. \tag{24}$$

Using (8), the elements of $\mathbf{R}_{Am}$ can be easily computed by the following equation:

$$R_{Am}(k, l) = \sum_{n=-N}^{N} b_s[n]r[(k-l)-n] + \sum_{n=k+l}^{N} b_s[n]r_c[(k+l)-n], \quad 1 \le k, \quad l \le v. \tag{25}$$

Finally, the sensitivity matrix for LSF parameters, $\mathbf{D}_w(\mathbf{w})$, based on the WLSD measure can be determined by (18) and the optimal WMSE weights, named OWMSE weights, are the diagonal elements of the sensitivity matrix $\mathbf{D}_w(\mathbf{w})$.

The computational complexity for calculating the OWMSE weights are evaluated as follows. The impulse response $h[n]$, for $0 \le n \le L$, requires vL multiply-adds (multiply-add meaning a multiplication operation followed by an addition operation), where v denotes the order of LPC model, L denotes the maximum time index of impulse response $h[n]$. Equations (23) and (24) require $[(L+1)-(N+v-1)/2](N+v)$ multiply-adds and $N(N+1)/2$ multiply-adds, respectively, where N denotes the length of $b_s[n]$ for $n > 0$. The first term and the second term in the left-hand side of (25) require $(2N+1)v$ multiply-adds and $(N-v)(2v-1)$ multiply-adds, respectively. The approach for computing the Jacobian matrix $\mathbf{J}_w(\mathbf{w})$ in [4] requires about $20v$ multiply-adds and v sine (sine function) and v cosine (cosine function) of LSF parameters. However, the cosine function of the LSF parameters is typically computed by the root-search procedure in finding LSF parameters. Finally, the computation for the diagonal terms of $\mathbf{D}_w(\mathbf{w})$ in (18) requires about (v^3+v^2) multiply-adds. Hence, the total computation complexity is

$$\mathbf{C}_{\text{OWMSE}} = \left\{ vL + \left[(L+1) - \frac{N+v-1}{2} \right] (N+v) + \frac{N(N+1)}{2} + (2N+1)v + (N-v)(2v-1) + 20v + v^3 + v^2 \right\} \text{ multiply-adds} + \text{v sine}. \tag{26}$$

III. Experimental Results

The speech database used in our experiments consists of Chinese sentences recorded from 26 speakers including 16 men and ten women. The database is divided into two parts. One is named training database including 17 speakers (ten men and seven women) and the other is named testing database including nine speakers (six men and three women) independent of those used in the training database. Speech signals are low-pass filtered at 4 kHz and digitized at a sampling rate of 8 kHz. A tenth-order LPC analysis based on Levinson–Durbin algorithm is performed every 20 ms using a 25 ms-length analysis Hamming window. Moreover, a fixed 15 Hz bandwidth expansion is applied to each pole of the LPC vector, by replacing a_i by $a_i\gamma^i$, for $1 \le i \le 10$, where $\gamma = 0.994$. The LPC parameters with bandwidth expansion of each frame are transformed to LSF parameters by using Chebyshev polynomials [13]. Notice that the spectra corresponding to silence are excluded. Finally, we have 95 753 LSF vectors for training in the training database and 60 366 LSF vectors for testing in the testing database.

TABLE I
Bit Allocation of each Stage of 4-Stage VQ for Various Bit Rates in the Experiments

Bit rates	Stage 1 (bits)	Stage 2 (bits)	Stage 3 (bits)	Stage 4 (bits)
24 bits	6	6	6	6
27 bits	7	7	7	6
30 bits	8	8	7	7

Since computing the fully quadratic distortion measure for each vector in a codebook during the LSF quantization is too complex, the simple WMSE distortion measure is used in our experiments. The codebooks are trained according to a multistage VQ structure by using the joint optimization procedure with four stages and search depth of $M = 8$ as described in [6]. The iterations before the error converges require about 100 time steps. The bit allocation of each stage for various bit rates is shown in Table I.

A. Experiment 1

Since balanced spectral error is desirable, in this experiment, we incorporate auditory perception into the LSF quantization by using the Bark weighting function [7]

$$W_B(w) = \frac{1}{25 + 75\left(1 + 1.4\left(\frac{F_s w}{2000\pi}\right)^2\right)^{0.69}} \tag{27}$$

where F_s is the sampling frequency of 8 kHz. Using the weighting function $W_B(w)$ and the results of previous section, the OWMSE weights are obtained. Here, a 256-point IDFT is applied on $W_B(w)$ to obtain the time sequence $\tilde{W}_b[n]$ with time indices from -127 to $+128$. Then, the impulse response $W_b[n]$ of the weighting function can be approximately obtained from $\tilde{W}_b[n]$ by setting $\tilde{W}_b[n]$ to zero if n is larger than 20 or less than -20, because the ratios of $\tilde{W}_b[n]$ to $\tilde{W}_b[0]$ for $n > 20$ or $n < -20$ are too small. Hence, the order of the impulse response $W_b[n]$ of the weighting function $W_B(w)$ used in this experiment is 40.

To compare the performance of the obtained OWMSE weights with that of the weights previously proposed by others, the LSF vector quantizers are also trained by using three other sets of weights. In [4], the optimal weights based on the LSD measure for LSF quantization were derived, named Gardner weights. In [7], a set of weights in which the weight for the ith LSF component is the product of the Gardner weight, G_i, and $W_B^2(w_i)$ in (27) were suggested, named Cohn weights. In [5], a set of weights were given subjectively based on the linear prediction power spectra, $P(w_i)$, in (2) as $w_i(w) = c_i[P(w_i)^{0.15}]$, where $c_i = 1$ for $1 \le i \le 8$, $c_9 = 0.8^2$, $c_{10} = 0.4^2$, named Paliwal weights.

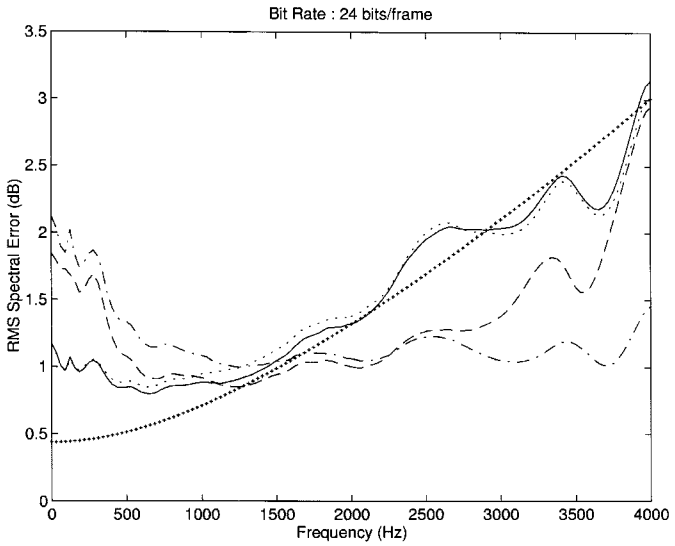

Fig. 1. Spectral error distributions for 24-bit LSF multistage vector quantizers using four different sets of weights in Experiment 1, where solid line corresponds to OWMSE weights, dotted line to Cohn weights, dashed line to Paliwal weights, dashdot line to Gardner weights, and "+" line corresponds to the inverse of the Bark weighting function multiplied by $4.387\,95 \times 10^{-3}$.

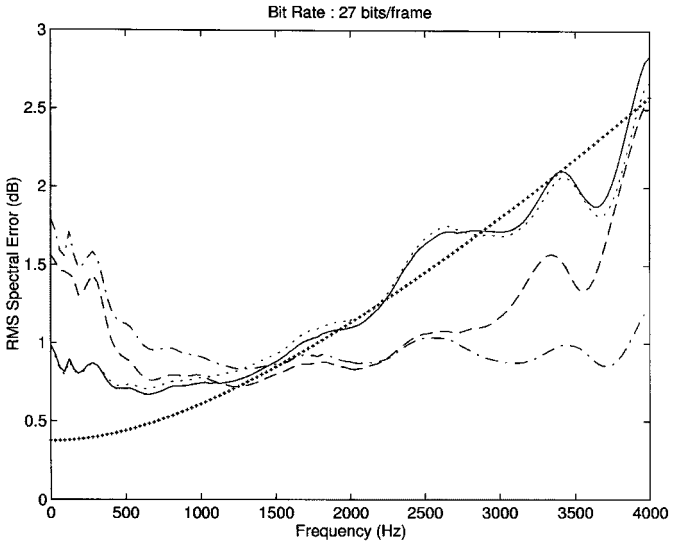

Fig. 2. Spectral error distributions for 27-bit LSF multistage vector quantizers using four different sets of weights in Experiment 1, where solid line corresponds to OWMSE weights, dotted line to Cohn weights, dashed line to Paliwal weights, dashdot line to Gardner weights, and "+" line corresponds to the inverse of the Bark weighting function multiplied by $3.760\,92 \times 10^{-3}$.

In order to observe the spectral error distribution, the RMS distortion is calculated by evaluating the original and quantized spectra given by (2) for 256 discrete frequencies, converting the distortion to dB at each frequency, and then taking the RMS value of the distortion at each frequency over all measured spectra. The experimentally obtained spectral error distributions at different bit rates are shown in Figs. 1–3. In order to contrast the shapes of the spectral error distribution curves in Figs. 1–3 with the shape of the inverse of the Bark weighting curve in (27), three curves of the inverse of the Bark weighting functions multiplied by $4.387\,95 \times 10^{-3}$, $3.760\,92 \times 10^{-3}$ and $3.082\,56 \times 10^{-3}$ are also plotted in Figs. 1–3, respectively. The constant gains of $4.387\,95 \times 10^{-3}$, $3.760\,92 \times 10^{-3}$ and $3.082\,56 \times 10^{-3}$ are obtained by minimizing the mean square error between the spectral error distribution curve for the OWMSE weights and the inverse of the Bark weighting curve. From Figs. 1–3, it is obvious that the spectral error distribution curve for the OWMSE weights is closer to the spectral error distribution of the inverse of the Bark weighting curve than those for the Paliwal weights and Gardner weights. Although the OWMSE weights and Cohn weights have similar spectral error distribution curves in the figures, the OWMSE weights still outperform the Cohn weights as will be shown in the following tests.

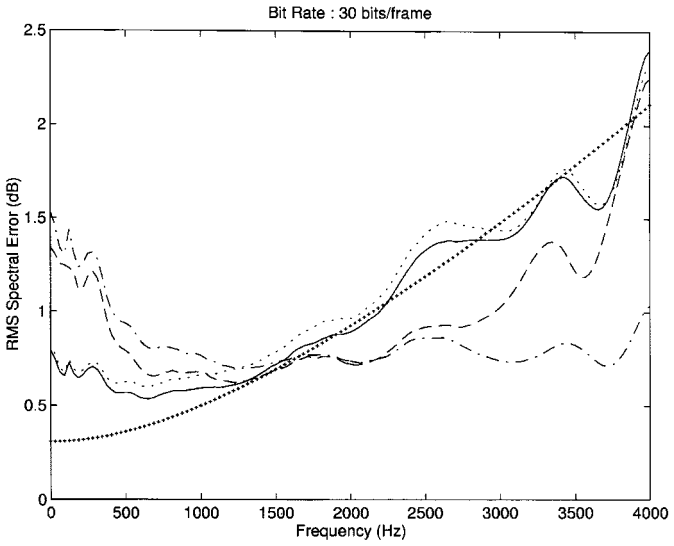

Fig. 3. Spectral error distributions for 30-bit LSF multistage vector quantizers using four different sets of weights in Experiment 1, where solid line corresponds to OWMSE weights, dotted line to Cohn weights, dashed line to Paliwal weights, dashdot line to Gardner weights, and "+" line corresponds to the inverse of the Bark weighting function multiplied by $3.082\,56 \times 10^{-3}$.

For further analyzing the experimental data and comparing the performance of the OWMSE weights with that of other weights, two performance measures are used. One is the average WLSD in decibels for measuring the quantization performance. The average WLSD is defined by (28), shown at the bottom of the next page, where

$W_{B0} = 2.603 \times 10^{-5}$	normalization factor of $W_B(w)$,
$S_n(w)$ and $\hat{S}_n(w)$	power spectra of the nth speech frame without and with quantization, respectively;
N_f	total count of frames in the testing database.

The other measure is the error balanced degree, which is defined as

$$\mathrm{EBD} = \min_{\kappa} \int_0^{\pi} (\overline{x}(w) - \kappa x(w))^2 \, dw \tag{29}$$

where $\overline{x}(w)$ and $x(w)$ represent the spectral error distribution curve and the inverse of the Bark weighting function, respec-

tively, and κ is a gain factor. The spectral error distribution curve $\overline{x}(w)$ in (29) is defined as

$$\overline{x}(w) = \sqrt{\frac{\sum_{i=1}^{N_e}[P_i(w) - \hat{P}_i(w)]^2}{N_e}} \tag{30}$$

where $P_i(w)$ and $\hat{P}_i(w)$ represent the ith original power spectra and the ith quantized spectra in decibels, respectively, and N_e is the number of frames. The concept for defining (29) comes from the minimization of error by simultaneously optimizing the gain and shape of one curve or vector, such as the training of gain-shape codebooks in [15]. The gain factor κ in (29) is to shift the inverse of the Bark weighting function to obtain the minimum EBD value, while keeping the original envelop shape of the inverse of the Bark weighting function unchanged. Hence, the EBD measure is a good indicator for evaluating the error balanced degree between the spectral error distribution curve and the inverse of the Bark weighting function. In our experiment, a discrete type of (29) is used,

$$\text{EBD} = \min_{\kappa} \sum_{i=0}^{I} (\overline{x}(w_i) - \kappa x(w_i))^2 \tag{31}$$

where $w_1, w_2, \cdots, w_I$ are the set of discrete frequencies. The optimal (minimum error) gain value κ_{opt} can be obtained as follows by setting to zero the differentiation of (31) with respect to κ,

$$\kappa_{opt} = \frac{\sum_{i=0}^{I} x(w_i)\overline{x}(w_i)}{\sum_{i=0}^{I} x^2(w_i)}. \tag{32}$$

Thus, by substituting (32) into (31), we can obtain

$$\text{EBD} = \frac{\sum_{i=0}^{I} x^2(w_i) \sum_{i=0}^{I} \overline{x}^2(w_i) - \left[\sum_{i=0}^{I} x(w_i)\overline{x}(w_i)\right]^2}{\sum_{i=0}^{I} x^2(w_i)} \tag{33}$$

for computing EBD values. In our experiment, I is set as 128 since 256-point discrete frequencies are used for computing the spectral error.

The values of $\overline{\text{WSD}}$ and EBD are calculated at different bit rates as listed in Tables II and III, respectively. These tables indicate that the values of $\overline{\text{WSD}}$ and EBD for our OWMSE weights are smaller than those for the other three sets of weights at the same bit rate. This shows the OWMSE weights obtained by our scheme outperform the compared counterparts. More precisely, the quantization error distortion distribution for the OWMSE weights is more balanced than that for other three sets of weights. Figs. 1–3 also reflect this phenomenon.

TABLE II
($\overline{\text{WSD}}$) VALUES OF THE LSF MULTISTAGE VECTOR QUANTIZERS USING OWMSE WEIGHTS, COHN WEIGHTS, PALIWAL WEIGHTS AND GARDNER WEIGHTS AT THREE DIFFERENT BIT RATES IN EXPERIMENT 1

	$\overline{\text{WSD}}$ in dB			
Bit rates	OWMSE weights	Cohn weights	Paliwal weights	Gardner weights
24	1.048	1.070	1.231	1.372
27	0.878	0.890	1.037	1.151
30	0.717	0.751	0.889	0.962

It is interesting to check how much performance is lost by the WMSE approximation since the error distortion measure with the spectral weighting function, $B(w)$, no longer results in a diagonal sensitivity matrix for LSF parameters according to Theorem 3. An experiment to check this was conducted for the bit rates shown in Table I. In this experiment, the codebooks are also trained using the full quadratic distortion measure according to the multistage VQ structure with the same optimization procedure, stages and search depth as described in the above experiment. The weighting function used for computing the LSF sensitivity matrix in this experiment is also the Bark weighting function as defined in (27). The average WLSD, $\overline{\text{WSD}}$, defined in (28) is used to observe the performance of VQs trained and quantized using the full quadratic distortion measure and the WMSE distortion measure. In the experiment, the values of $\overline{\text{WSD}}$ for VQs using the full quadratic distortion measure at 30, 27, and 24 bits/frame are 0.708 dB, 0.871 dB and 1.044 dB, respectively. Since the difference at each bit rate between the value of $\overline{\text{WSD}}$ for VQ using the full quadratic distortion measure and that, listed in Table II, for VQ using the WMSE distortion measure is small, we can see that the performance lost by the WMSE approximation is not significant. This experimental result also shows that the nondiagonal terms in the sensitive matrix for LSF parameters using Bark weighting function at 30, 27, and 24 bits/frame maybe neglected, if the computation complexity is the major concern. It is worth pointing out that given the multistage nature of the VQ scheme, the full quadratic distortion measure can be used only in the final stage of the quantizer. This can decrease the computation required while still provide the *optimal* weighting in the final stage, where the high-rate approximations are most likely to be valid and valuable. Obviously, the performance (i.e., the average WLSD values) of this scheme will lie between those of the two extreme cases, one for the WMSE distortion measure in each stage, the other for the full quadratic distortion measure in each stage.

$$\overline{\text{WSD}} = \frac{1}{N_f} \sum_{n=1}^{N_f} \sqrt{\frac{1}{W_{B_0}\pi} \int_0^{\pi} W_B^2(w)[10\log_{10}(S_n(w) - 10\log_{10}\hat{S}_n(w))]^2\, dw} \tag{28}$$

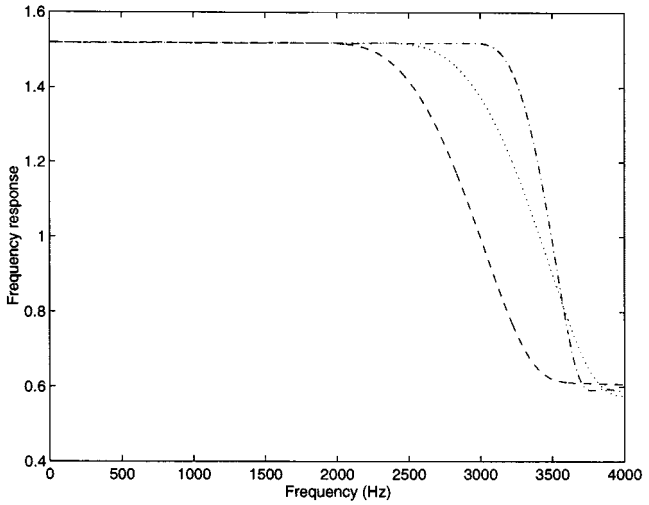

Fig. 4. Frequency responses of the three weighting functions used in Experiment 2, where dashed line corresponds to the weighting function h_{l1}, dotted line to the weighting function h_{l2}, dashdot line to the weighting function h_{l3}.

Based on the above experimental results and to avoid the computation complexity with full quadratic distortion measure for each input vector of quantizers during a codebook search procedure in LSF quantization, the simple WMSE distortion measure is used in all the other experiments in the rest of this section. It can be expected that better performance can be obtained if the full quadratic distortion measure is used, instead.

B. Experiment 2

Paliwal and Atal in [5] observed that the transparent quantization of LPC information can be obtained if the following three conditions are met: 1) the average spectral distortion (SD) is about 1 dB; 2) there is no outlier frame having SD larger than 4 dB, and 3) the number of outlier frames having SD in the range of 2–4 dB is less than 2%. Here, SD is the same as D defined in (1) except that the integration range is in 0–3 kHz. According to the above conditions, since only the lower 3/4 of the frequency range is used for computing the spectral distortion error, we can design a weighting function which has larger function values at the lower 3/4 of the frequency range, and much smaller function values at higher frequency range to reduce the quantization SD.

Three weighting functions used in this experiment are shown in Fig. 4. All of the three weighting functions have larger values at lower frequencies than those at higher frequencies, and are designed as FIR filters for easily computing the OWMSE weights as mentioned in Section II-C. According to the different bandwidths of the three weighting functions, we denote the weighting function with the narrowest bandwidth as $h_{l1}(w)$, the one with the widest bandwidth as $h_{l3}(w)$, and the other as $h_{l2}(w)$. The orders of the three weighting functions, $h_{l1}(w)$, $h_{l2}(w)$ and $h_{l3}(w)$, are 20, 20, and 40, respectively. The OWMSE weights corresponding to the three weighting functions, $h_{l1}(w)$, $h_{l2}(w)$ and $h_{l3}(w)$, obtained from the procedure in Section II-C are denoted by h_{l1} weights, h_{l2} weights, and h_{l3} weights, respectively. In order to test the performance of LSF quantization using these OWMSE weights at different bit rates, the results of LSF quantization using Paliwal weights at the same bit rates are also obtained for comparisons. The definition of Paliwal weights has been given in Experiment 1.

TABLE III
(EBD) VALUES OF THE LSF MULTISTAGE VECTOR QUANTIZERS USING OWMSE WEIGHTS, COHN WEIGHTS, PALIWAL WEIGHTS AND GARDNER WEIGHTS AT THREE DIFFERENT BIT RATES IN EXPERIMENT 1

EBD				
Bit rates	OWMSE weights	Cohn weights	Paliwal weights	Gardner weights
24	8.626	10.779	32.127	60.549
27	5.645	6.729	23.082	42.866
30	3.600	4.660	16.917	30.271

TABLE IV
($\overline{\text{WSD}}$) VALUES OF THE LSF MULTISTAGE VECTOR USING PALIWAL WEIGHTS AND OWMSE WEIGHTS WITH WEIGHTING FUNCTIONS $h_{l1}(w)$, $h_{l2}(w)$, $h_{l3}(w)$, AT THREE DIFFERENT BIT RATES IN EXPERIMENT 2

WSD in dB						
Weighting function	$h_{l1}(w)$		$h_{l2}(w)$		$h_{l3}(w)$	
Bit rates	OWMSE	Paliwal	OWMSE	Paliwal	OWMSE	Paliwal
24	1.472	1.525	1.617	1.630	1.676	1.692
27	1.241	1.283	1.355	1.373	1.382	1.427
30	1.031	1.100	1.142	1.180	1.167	1.228

TABLE V
EBD VALUES OF THE LSF MULTISTAGE VECTOR QUANTIZERS USING PALIWAL WEIGHTS AND OWMSE WEIGHTS WITH WEIGHTING FUNCTIONS $h_{l1}(w)$, $h_{l2}(w)$, $h_{l3}(w)$, AT THREE DIFFERENT BIT RATES IN EXPERIMENT 2

EBD						
Weighting function	$h_{l1}(w)$		$h_{l2}(w)$		$h_{l3}(w)$	
Bit rates	OWMSE	Paliwal	OWMSE	Paliwal	OWMSE	Paliwal
24	43.349	58.036	53.574	56.273	69.901	70.051
27	29.198	40.756	36.895	39.369	48.218	49.496
30	19.836	29.960	25.636	29.213	34.566	37.015

Similar to Tables II and III in Experiment 1, Tables IV and V show that the performance comparisons between the vector quantizer using the Paliwal weights and the vector quantizer using the OWMSE weights with different weighting functions $h_{l1}(w)$, $h_{l2}(w)$, $h_{l3}(w)$, where the error index $\overline{\text{WLSD}}$ and EBD are defined in Eqs. (28) and (29), respectively, except that the term $W_B(w)$ in (28) is replaced by one of $h_{l1}(w)$, $h_{l2}(w)$, and $h_{l3}(w)$ here. From Tables IV and V, we can see that the performance of the vector quantizers using the OWMSE weights with either of the three weighting functions is better than the vector quantizer using the Paliwal weights. In addition, to observe the spectral error distributions at different bit rates, we show the spectral error distribution curves in Figs. 5–7. In Figs. 5–7, we observe that the OWMSE weights can be efficiently used for shaping the quantization spectral error distribution into any desired frequency-dependent curve depending on the selected weighting functions at different bit rates. Finally, to check the performance in the sense of *transparent quantization* of LPC information, we list the average SD values of the vector quantizers using the Paliwal weights or the OWMSE weights with different weighting functions in Table VI. Table VI shows that the performance of LSF quantization using the OWMSE weights is better than that using the Paliwal weights even at lower bit rate, 24 bits/frame. This conclusion comes from the observation on Table VI that the average SD values using the OWMSE weights are smaller than those using the Paliwal weights, and the outliers within 2–4 dB and above 4 dB using

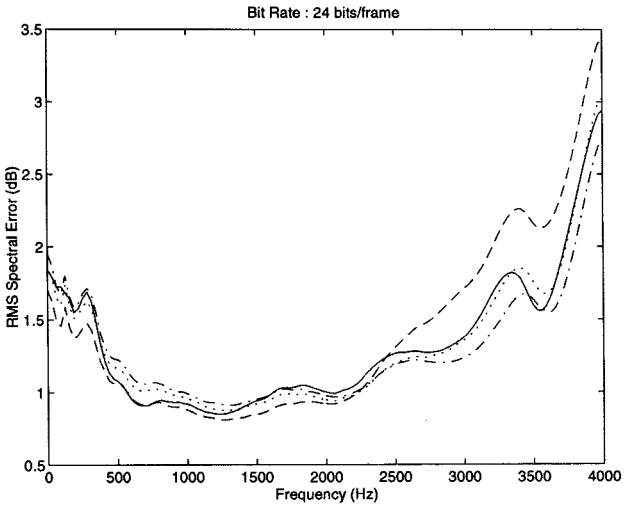

Fig. 5. Spectral error distributions for 24-bit LSF multistage vector quantizers using four different sets of weights in Experiment 2, where solid line corresponds to Paliwal weights, dashed line to h_{l1} weights, dotted line to h_{l2} weights, dashdot line to h_{l3} weights.

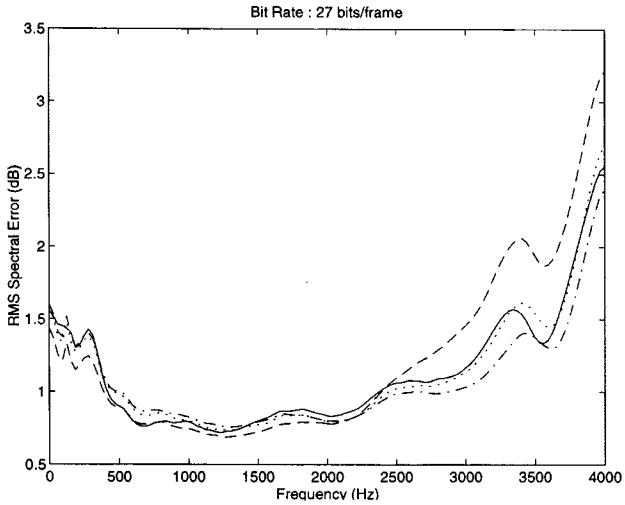

Fig. 6. Spectral error distributions for 27-bit LSF multistage vector quantizers using four different sets of weights in Experiment 2, where solid line corresponds to Paliwal weights, dashed line to h_{l1} weights, dotted line to h_{l2} weights, dashdot line to h_{l3} weights.

the OWMSE weights are comparable to those using the Paliwal weights.

From Table VI, we can see that the average SD values using the weighting function $h_{l3}(w)$ are somewhat larger than those using the other weighting functions at the same bit rate. This phenomenon is explained as follows. Since the values of the weighting function $h_{l3}(w)$ are larger than those of the others in almost all the higher frequency range above 3 kHz, the quantization errors using the weighting function $h_{l3}(w)$ are smaller than those using the others in the higher frequency range above 3k Hz as shown in Figs. 5–7. However, it is a fact that the total quantization errors at the same bit rate are fixed regardless of what kind of error shaping function being used in the quantizers. Therefore, it is reasonable that the average SD values using the weighting function $h_{l3}(w)$ are higher than those using the others, since the average SD values only contain the quantization errors in the lower frequency range of 0–3 kHz.

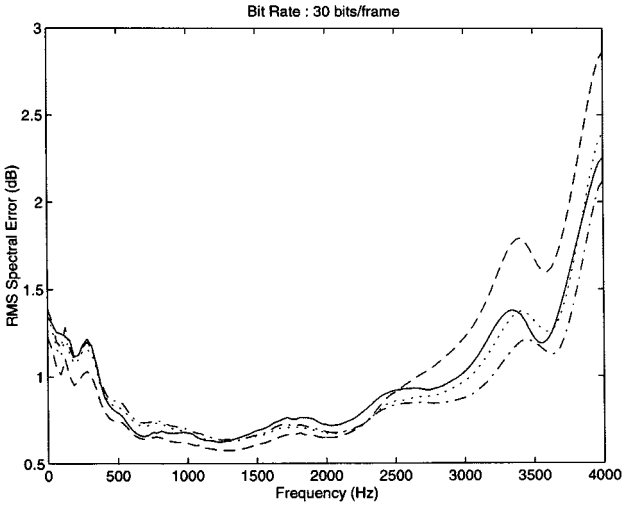

Fig. 7. Spectral error distributions for 30-bit LSF multistage vector quantizers using four different sets of weights in Experiment 2, where solid line corresponds to Paliwal weights, dashed line to h_{l1} weights, dotted line to h_{l2} weights, dashdot line to h_{l3} weights.

TABLE VI
SD VALUES OF THE LSF MULTISTAGE VECTOR QUANTIZERS USING PALIWAL WEIGHTS, h_{l1} WEIGHTS, h_{l2} WEIGHTS, AND h_{l3} WEIGHTS AT THREE DIFFERENT BIT RATES IN EXPERIMENT 2

Bit rates	SD computed in 0 - 3kHz band			
	Weights	Avg. SD (dB)	Outliers (%)	
			2-4 dB	> 4 dB
24	Paliwal	1.105	2.730	0.010
	h_{l1}	1.087	2.564	0.010
	h_{l2}	1.095	2.322	0.013
	h_{l3}	1.098	2.791	0.008
27	Paliwal	0.928	1.052	0.003
	h_{l1}	0.911	1.024	0.003
	h_{l2}	0.915	0.977	0.007
	h_{l3}	0.919	0.837	0.007
30	Paliwal	0.796	0.480	0.003
	h_{l1}	0.755	0.375	0.007
	h_{l2}	0.770	0.393	0.003
	h_{l3}	0.776	0.396	0.007

TABLE VII
COMPUTATIONAL COMPLEXITY OF THE ALGORITHMS IN CALCULATING THE OWMSE WEIGHTS, THE COHN WEIGHTS, THE GARDNER WEIGHTS, THE PALIWAL WEIGHTS, AND THE SIMPLE MSE MEASURE

Weighting types	Computational complexity
OWMSE	$\mathbf{C}_{\text{OWMSE}}$ in Eq. (26)
Cohn	$\{vL + [(L+1) - v/2]v + 20v + v^2(v+1)/2\}$ multiply-adds + v sine + v division + v power + $3v$ multiply-adds
Gardner	$\{vL + [(L+1) - v/2]v + 20v + v^2(v+1)/2\}$ multiply-adds + v sine
Paliwal	$12v$ multiply-adds + $10v$ cosine + $10v$ sine + v division + v power
Simple MSE	0 (No weighting)

The comparisons of the computational complexity required for computing the OWMSE weights to that required for computing the Cohn weights, the Gardner weights, the Paliwal weights, and a simple MSE measure are listed in Table VII, assuming that the autocorrelation of the LPC impulse response $r[k]$ is computed by (23). In this table, L denotes the maximum time index of the impulse response of LPC model, N denotes the length of the impulse response $b_s[n]$ (for $n > 0$) of the weighting function $B_s(w)$, v denotes the order of LPC model,

"multiply-add" denotes the multiplication operation followed by an addition operation, "division" denotes the division operation, "sine" denotes the operation of the sine function, "cosine" denotes the operation of the cosine function, and "power" denotes the operation of the power function, where sine and cosine functions are typically computed by some numerical method such as the Taylor series expansion. In Table VII, the term "$v^2(v+1)/2$ multiply-adds" is the requirement for efficiently computing the diagonal terms of the sensitivity matrix for LSF parameters in [4]. In computing the Cohn weights, the additional computation requirement, v division $+v$ power $+3v$ multiply-adds, comparing to the calculation of the Gardner weights, comes from the computation of the Bark weighting function in (27) at the v LSF parameters. From Table VII, we can see that the OWMSE weighting requires more multiply-adds than the other weightings but it requires no divisions and power operations. For better understanding, the numerical values of the computational complexity for the case of $N = 20$, $L = 400$, and $v = 10$, as in our Experiment 1 are found to be 17 885, 10 390, 970, 4580, and 0 multiply-adds for the OWMSE weights, the Cohn weights, the Gardner weights, the Paliwal weights, and the simple MSE measure, respectively, according to the equations in Table VII, where the computational complexities for "division," "sine," "cosine," and "power" in Table VII require about 18, 18, 20, and 48 multiply-adds estimated according to the runtime support functions of the TMS320C30 DSP chip. In this case, the required DSP MIPS for the OWMSE weights, and the four compared schemes are 0.894, 0.520, 0.485, 0.229 and 0, respectively, where the estimation is based on the TMS320C30 DSP chip produced by TI (Texas Instrument, Inc.).

C. Experiment 3

The results of Experiments 1 and 2 show that the LSF quantization using weights do obtain the desired effect of weighting the spectral error toward the low frequencies. To further test how this affects the overall speech quality in an actual speech coding algorithm, we add weighted spectral error into LSF quantization in a CELP coder. In this experiment, the codebooks trained at 30, 27, and 24 bits/frame with the OWMSE weights obtained in Experiment 1 (for the Bark weighting function) and Experiment 2 (for the weighting function h_{l1}) are used for this testing. Three speech samples are prepared for coding; each of them is analyzed and then reconstructed by the original CELP coder and the modified CELP coders with weighted spectral error incorporated into LSF quantizer obtained in Experiment 1 and Experiment 2. The LSF quantizer used in the original CELP coder is a nonuniform scalar quantizer using 34 bits. This experiment produces 21 reconstructed speech signals; three for the original CELP coder, nine (3 speech signals by 3 different bit rates) for the modified CELP coders corresponding to the Bark weighting function in Experiment 1, and nine for the modified CELP coders corresponding to the weighting function, h_{l1}, in Experiment 2. The three speech samples and the 21 reconstructed speech signals can be found in the web site: http://falcon3.cn.nctu.edu.tw/~ncw/shaping. In an informal test, we find that the speech quality of the reconstructed speech signals using the modified CELP coders are comparable to that using the original CELP coder even at lower bit rates such as 24 bits/frame. In other words, the results show that the approach incorporating the weighted spectral error in LSF quantization into a CELP coder does not produce strange distortion, and moreover, results in lower bit rates for equivalent speech quality, even at 24 bits/frame.

IV. Conclusions

This paper has presented an error shaping technique of LSF vector quantization (VQ) based on the WLSD measure. It has been shown that the VQ trained by the quadratically weighted measure converges to a VQ trained by the WLSD measure. The quadratic weighting matrix, the so-called "sensitivity matrix," is given by the second term of Taylor series expansion of the WLSD measure and has been proven to be a nondiagonal matrix. The approximate computation algorithm for calculating the quadratic weighting matrix is provided. The optimal WMSE weights of LSF quantization are determined by the diagonal elements of the quadratic weighting matrix. The proposed error shaping technique has been applied to make better use of the human perceptual characteristics based on the WMSE measure. For the balanced spectral quantization error in Experiment 1, it has been shown that the performance of the LSF quantization using the optimal WMSE weights is better than those using other previously proposed weights, including Paliwal weights, Gardner weights and Cohn weights. For the transparent spectral error quantization in Experiment 2, it has been shown that the performance of the LSF quantization using the proposed optimal WMSE weights is better than that using Paliwal weights. Finally, the test of incorporating the proposed error shaping technique into the LSF quantization (34 bits/frame originally) of an actual CELP coder in Experiment 3 has shown that the proposed scheme can result in lower bit rates (e.g., 24 bits/frame) for equivalent speech quality.

Appendix A

This appendix provides the proof of Theorem 2.

Let $a_0 = \hat{a}_0 = -1$. Then, we have $A(w) = -\sum_{m=0}^{v} a_m e^{-jwm}$ and $\hat{A}(w) = -\sum_{m=0}^{v} \hat{a}_m e^{-jwm}$. This also results in $|A(w)|^2 = \sum_{m=0}^{v}\sum_{n=0}^{v} a_m a_n \cos(w(m-n))$. By substituting these results in (6) and performing simple differentiation [4], we can show that

$$\begin{aligned}
&\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{a},\hat{\mathbf{a}})}{\partial \hat{a}_k \partial \hat{a}_l}\right|_{\hat{\mathbf{a}}=\mathbf{a}} \\
&\quad = 4\alpha \sum_{m=0}^{v}\sum_{n=0}^{v} a_m a_n \frac{1}{2\pi}\int_{-\pi}^{\pi} B^2(w) \\
&\qquad \cdot \frac{\cos(w(k+l-m-n)) + \cos(w(k-l-m+n))}{|A(w)|^4}\, dw \\
&\quad = 4\alpha \sum_{m=0}^{v}\sum_{n=0}^{v} a_m a_n [f(k+l-m-n) + f(k-l-m+n)],
\end{aligned} \tag{A.1}$$

where

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} B_s(w) \frac{e^{jwx}}{|A(w)|^4}\, dw \tag{A.2}$$

is the inverse discrete-time Fourier transform of $B_s(w)/|A(w)|^4$, since $B_s(w)$ is a real function. Thus we have $F(w) = B_s(w)/|A(w)|^4$, and $F(z) = B_s(z)/(A^2(z)A^2(z^{-1}))$.

Using the same notations in [4], let

$$M(x) = \sum_{m=0}^{v}\sum_{n=0}^{v} a_m a_n f(x-m-n) \tag{A.3}$$

and

$$N(x) = \sum_{m=0}^{v}\sum_{n=0}^{v} a_m a_n f(x-m+n). \tag{A.4}$$

Then we have

$$\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{a},\hat{\mathbf{a}})}{\partial \hat{a}_k \partial \hat{a}_l}\right|_{\hat{\mathbf{a}}=\mathbf{a}} = 4\alpha(M(k+l)+N(k-l)). \tag{A.5}$$

Since

$$M(z) = A^2(z)F(z) = \frac{B_s(z)}{A^2(z^{-1})} \tag{A.6}$$

and

$$N(z) = A(z)A(z^{-1})F(z) = \frac{B_s(z)}{A(z)A^(z^{-1})} \tag{A.7}$$

we can take inverse Z transforms of $M(z)$ and $N(z)$, and use the Z transform pairs of Eqs. (9) and (10) to obtain the following equality

$$\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{a},\hat{\mathbf{a}})}{\partial \hat{a}_k \partial \hat{a}_l}\right|_{\hat{\mathbf{a}}=\mathbf{a}} = 4\alpha \sum_{n=-\infty}^{\infty} b_s[n](r_c[(k+l)-n]+r[(k-l)-n]). \tag{A.8}$$

This completes the proof. □

APPENDIX B

This appendix provides the proof of Theorem 3.

Before going to the proof of Theorem 3, some notations are defined first. Let LSF parameters be denoted by $\mathbf{w} = [w_1, w_2, \cdots, w_v]^T$. Since the roots of $P(z)$ correspond to the odd indices of LSF parameters and the roots of $Q(z)$ correspond to the even indices of LSF parameters, $A(w)$ can be rewritten as

$$A(w) = \tfrac{1}{2}(P(w)+Q(w)) \tag{B.1}$$

where

$$P(w) = (1+e^{-jw}) \prod_{i\ \mathrm{odd}} (1-2\cos w_i e^{-jw}+e^{-2jw}) \tag{B.2}$$

$$Q(w) = (1-e^{-jw}) \prod_{i\ \mathrm{even}} (1-2\cos w_i e^{-jw}+e^{-2jw}). \tag{B.3}$$

For notational clarity, we define the following equations:

$$\begin{aligned}
\overline{p}_0(w) &= 1+e^{-jw} \\
\overline{p}_i(w) &= 1-2\cos(w_{2i-1})e^{-jw}+e^{-2jw} \\
\overline{q}_0(w) &= 1-e^{-jw} \\
\overline{q}_i(w) &= 1-2\cos(w_{2i})e^{-jw}+e^{-2jw}.
\end{aligned} \tag{B.4}$$

Then we have

$$P(w) = \prod_{i=0}^{v/2} \overline{p}_i(w), \qquad Q(w) = \prod_{i=0}^{v/2} \overline{q}_i(w). \tag{B.5}$$

The simple differentiation of $A(w)$ with respect to the ith LSF parameter is denoted by [4]

$$\begin{aligned}
J_i(w) = \frac{\partial A(w)}{\partial w_i} &= \begin{cases} \dfrac{1}{2}\dfrac{\partial P(w)}{\partial w_i}; & i \quad \text{odd} \\ \dfrac{1}{2}\dfrac{\partial Q(w)}{\partial w_i}; & i \quad \text{even} \end{cases} \\
&= \begin{cases} \sin(w_i)e^{-jw} \displaystyle\prod_{j=0;\, j\neq(i+1)/2}^{v/2} \overline{p}_j(w); & i \quad \text{odd} \\ \sin(w_i)e^{-jw} \displaystyle\prod_{j=0;\, j\neq i/2}^{v/2} \overline{q}_j(w); & i \quad \text{even}. \end{cases}
\end{aligned} \tag{B.6}$$

With the above notation definitions, we can now proceed the proof of Theorem 3 as follows.

Proof of Theorem 3: Computing the elements of the sensitivity matrix for LSF parameters by simple differentiation in (6) directly, we obtain the following equation:

$$\begin{aligned}
&\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{w},\hat{\mathbf{w}})}{\partial \hat{w}_k \partial \hat{w}_l}\right|_{\hat{\mathbf{w}}=\mathbf{w}} \\
&= \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w) \frac{\dfrac{\partial A^*(w)}{\partial w_k}A(w)+A^*(w)\dfrac{\partial A(w)}{\partial w_k}}{|A(w)|^2} \\
&\quad \cdot \frac{\dfrac{\partial A^*(w)}{\partial w_l}A(w)+A^*(w)\dfrac{\partial A(w)}{\partial w_l}}{|A(w)|^2}\, dw
\end{aligned} \tag{B.7}$$

where "$*$" is the complex conjugate operator. The proof of Theorem 3 is completed if one nondiagonal element is proved to be nonzero when the length of impulse response of the weighting function used in the WLSD measure is larger than 3. Hence, we shall prove this theorem by showing that there exists at least one nondiagonal element in (B.7) (i.e., $k \neq l$) which is nonzero. In the following, we consider the case of odd k and even l, i.e., $\hat{w}_k$ and $\hat{w}_l$ in (B.7) are the roots of $P(z)$ and $Q(z)$, respectively.

By substituting (B.6) into (B.7), we obtain

$$\begin{aligned}
&\left.\frac{\partial^2 \mathrm{WLSD}(\mathbf{w},\hat{\mathbf{w}})}{\partial \hat{w}_k \partial \hat{w}_l}\right|_{\hat{\mathbf{w}}=\mathbf{w}} \\
&= \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w) \left[\frac{J_k^*(w)J_l^*(w)}{A^{*2}(w)}+\frac{J_k(w)J_l(w)}{A^2(w)}\right. \\
&\qquad \left.+\frac{J_k^*(w)J_l(w)}{A^*(w)A(w)}+\frac{J_k(w)J_l^*(w)}{A^*(w)A(w)}\right] dw.
\end{aligned} \tag{B.8}$$

Let

$$T(w) = \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w)\frac{J_k^*(w)J_l(w)}{A^*(w)A(w)}\, dw \tag{B.9}$$

and

$$U(w) = \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w)\frac{J_k(w)J_l^*(w)}{A^*(w)A(w)}\,dw. \qquad \text{(B.10)}$$

Then, using the symmetry properties of P polynomials, the antisymmetry properties of Q polynomials and symmetry property of $B_s(w)$, we can show

$$\begin{aligned} T(w) &= \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w) \\ &\quad \cdot \frac{(e^{j(v+1)w}J_k^*(-w))(-e^{-j(v+1)w}J_l(-w))}{|A(w)|^2}\,dw \\ &= -\frac{\alpha}{\pi}\int_{-\pi}^{\pi}\frac{B_s(-w)J_k^*(-w)J_l(-w)}{|A(-w)|^2}\,dw \\ &= -\frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w)\frac{J_k^*(w)J_l(w)}{A^*(w)A(w)}\,dw \\ &= -T(w). \end{aligned} \qquad \text{(B.11)}$$

Equation (B.11) results in $T(w) = 0$. Using the steps taken above, a similar proof can be derived for the fact $U(w) = 0$.

Now, substituting (B.6) into (B.8) again and using the above results, we can rewrite (B.8) as shown in (B.12) at the bottom of the page.

Let

$$G(w) = B_s(w)\frac{\left[\sin w_k \prod_{j\neq(k+1)/2}^{v/2}\overline{p}_j(w)\right]\left[\sin w_l \prod_{j\neq l/2}^{v/2}\overline{q}_j(w)\right]}{A^2(w)} \qquad \text{(B.13)}$$

and define the discrete-time Fourier pairs

$$\begin{aligned} g[n] &\Longleftrightarrow G(w), \\ \overline{p}_0[n] &= \delta[n] + \delta[n-1] \\ &\Longleftrightarrow \overline{p}_0(w), \\ \overline{p}_i[n] &= \delta[n] - 2\cos(w_{2i-1})\delta[n-1] + \delta[n-2] \\ &\Longleftrightarrow \overline{p}_i(w), \\ \overline{q}_0[n] &= \delta[n] - \delta[n-1] \\ &\Longleftrightarrow \overline{q}_0(w), \\ \overline{q}_i[n] &= \delta[n] - 2\cos(w_{2i})\delta[n-1] + \delta[n-2] \\ &\Longleftrightarrow \overline{q}_i(w) \end{aligned} \qquad \text{(B.14)}$$

where $g[n]$ is the convolution of $b_s[n]$ with the sequences $\overline{p}_n$ and $\overline{q}_m$ for all n and m not equal to $(k+1)/2$ and $l/2$, respectively. Then define

$$\begin{aligned} \tilde{H}[n] = &\overline{p}_0[n] \otimes \overline{p}_1[n] \otimes \cdots \overline{p}_i[n] \otimes \cdots \overline{p}_{v/2[n]} \\ &\otimes \overline{q}_0[n] \otimes \overline{q}_1[n] \otimes \cdots \overline{q}_j[n] \otimes \cdots \overline{q}_{v/2}[n] \\ &\otimes h[n] \otimes h[n], \qquad i \neq (k+1)/2, \qquad j \neq l/2 \end{aligned} \qquad \text{(B.15)}$$

where "$\otimes$" denotes the convolution operator. Since $B_s(w) = B_s^*(-w)$ holds and $\tilde{H}[n]$ is real and causal, $g[n]$ is real and noncausal. Without loss of generality, assume the length of the impulse response of $b_s[n]$ is $2N+1$ for $-N \le n \le N$, where N may be infinity. Then, it is easy to show that

$$\begin{aligned} g[-2] &= b_s[n] \otimes \tilde{H}[n]\Big|_{n=-2} \\ &= \begin{cases} \text{zero}, & \text{for } N < 2 \\ \displaystyle\sum_{i=2}^{N} b_s[i]\tilde{H}[i-2], & \text{for } N \ge 2 \end{cases} \end{aligned} \qquad \text{(B.16)}$$

where $\sum_{i=2}^{N} b_s[i]\tilde{H}[i-2]$ for $N \ge 2$ is always nonzero since one cannot choose a sequence $b_s[n]$ such that $\sum_{i=2}^{N} b_s[i]\tilde{H}[i-2] = 0$ for all time-varying signals $\tilde{H}[n]$. Now, (B.12) becomes

$$\begin{aligned} &\frac{\partial^2 \mathrm{WLSD}(\mathbf{w},\hat{\mathbf{w}})}{\partial\hat{w}_k\partial\hat{w}_l}\Big|_{\hat{\mathbf{w}}=\mathbf{w}} \\ &= \frac{2\alpha\sin w_k \sin w_l}{2\pi}\int_{-\pi}^{\pi}(G(w)e^{-j2w} + G^*(w)e^{j2w})\,dw \\ &= 2\alpha\sin w_k \sin w_l(g[n]|_{n=-2} + g^*[n]|_{n=-2}) \\ &= 4\alpha\sin w_k \sin w_l g[-2] \\ &= \begin{cases} \text{zero}, & \text{for } N < 2, \\ 4\alpha\sin w_k \sin w_l \displaystyle\sum_{i=2}^{N} b_s[i]\tilde{H}[i-2], & \text{for } N \ge 2 \end{cases} \end{aligned} \qquad \text{(B.17)}$$

where $4\alpha\sin w_k \sin w_l \sum_{i=2}^{N} b_s[i]\tilde{H}[i-2]$ for $N \ge 2$ is always nonzero due to the same reason as in (B.16).

This completes the proof. □

$$\begin{aligned} \frac{\partial^2 \mathrm{WLSD}(\mathbf{w},\hat{\mathbf{w}})}{\partial\hat{w}_k\partial\hat{w}_l}\Big|_{\hat{\mathbf{w}}=\mathbf{w}} &= \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w)\left[\frac{J_k^*(w)J_l^*(w)}{A^{*2}(w)} + \frac{J_k(w)J_l(w)}{A^2(w)}\right]dw. \\ &= \frac{\alpha}{\pi}\int_{-\pi}^{\pi} B_s(w)\frac{\left[\sin w_k e^{jw}\prod_{j\neq(k+1)/2}^{v/2}\overline{p}_j^*(w)\right]\left[\sin w_l e^{jw}\prod_{j\neq l/2}^{v/2}\overline{q}_j^*(w)\right]}{A^{*2}(w)} \\ &\quad + B_s(w)\frac{\left[\sin w_k e^{-jw}\prod_{j\neq(k+1)/2}^{v/2}\overline{p}_j(w)\right]\left[\sin w_l e^{-jw}\prod_{j\neq l/2}^{v/2}\overline{q}_j(w)\right]}{A^2(w)}\,dw. \end{aligned} \qquad \text{(B.12)}$$

REFERENCES

[1] F. Itakura, "Line spectrum representation of linear predictive coefficients of speech siganls," *J. Acoust. Soc. Amer.*, vol. 57, p. 35, Apr. 1975.

[2] F. K. Soong and B. H. Juang, "Line spectrum pair (LSP) and speech data compression," in *Proc. Int. Conf. Acoustic, Speech, Signal Processing '84*, 1984, pp. 1.10.1–1.10.4.

[3] A. McCree, K. Truong, E. B. George, T. P. Barnwell, and V. Viswanathan, "A 2.4 Kbit/s MELP coder candidate for new U.S. federal standard," in *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Processing '96*, pp. 200–203.

[4] W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," *IEEE Trans. Speech Audio Processing*, vol. 3, pp. 367–381, Sept. 1995.

[5] K. K. Paliwal and B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech Audio Processing*, vol. 1, pp. 3–14, Jan. 1993.

[6] W. P. LeBlanc, B. Bhattacharya, S. A. Mahmoud, and V. Cuperman, "Efficient search and design procedures for robust multistage VQ of LPC parameters for 4 kb/s speech coding," *IEEE Trans. Speech Audio Processing*, vol. 1, pp. 373–385, Oct. 1993.

[7] R. P. Cohn and J. S. Collura, "Incorporating perception into LSF quantization—Some experiments," in *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Processing '97*, pp. 1347–1350.

[8] A. H. Gray Jr. and J. D. Markel, "Distance measures for speech processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 380–391, Oct. 1976.

[9] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1082–1091, May 1999.

[10] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, pp. 329–330.

[11] J. S. Collura, A. V. McCree, and T. E. Tremain, "Perceptually based distortion measurements for spectrum quantization," in *Proc. 1995 IEEE Workshop Speech Coding Telecommunications*, Annapolis, MD, pp. 49–50.

[12] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 129–137, 1982.

[13] P. Kabal and P. Ramachadran, "The computation of line spectral frequencies using Chebyshev polynomials," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP-34, pp. 1419–1426, Dec. 1986.

[14] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*. Berlin, Germany: Springer-Verlag, 1976, pp. 31–32.

[15] A. M. Kondoz, *Digital Speech Coding for Low Bit Rate Communications Systems*. New York: Wiley, 1995, pp. 29–30.

Hsi-Wen Nein received the B.S. degree in control engineering from the National Chiao-Tung University, Hsinchu, Taiwan, R.O.C., in 1995, where he is currently pursuing the Ph.D. degree in the Department of Electrical and Control Engineering.

His current research interests include speech coding, speech synthesis, speech recognition, neural networks, and fuzzy systems.

Chin-Teng Lin (S'88–M'91–SM'99) received the B.S. degree in control engineering from the National Chiao-Tung University (NCTU), Hsinchu, Taiwan, R.O.C., in 1986 and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1989 and 1992, respectively.

Since August 1992, he has been with the College of Electrical Engineering and Computer Science, NCTU, where he is currently a Professor and Chairman of electrical and control engineering. He served as the Deputy Dean of the Research and Development Office, NCTU, from 1998 to 2000. His current research interests are fuzzy systems, neural networks, intelligent control, human–machine interface, and video and audio processing. He is the coauthor of *Neural Fuzzy Systems—A Neuro-Fuzzy Synergism to Intelligent Systems* (Englewood Cliffs, NJ: Prentice Hall) and the author of *Neural Fuzzy Control Systems with Structure and Parameter Learning* (Amsterdam, The Netherlands: World Scientific). He has published about 60 journal papers in the areas of neural networks and fuzzy systems.

Dr. Lin is a member of Tau Beta Pi and Eta Kappa Nu. He is also a member of the IEEE Computer Society, the IEEE Robotics and Automation Society, and the IEEE Systems, Man, and Cybernetics Society. He has been the Executive Council Member of Chinese Fuzzy System Association (CFSA) since 1995, and the Supervisor of Chinese Automation Association since 1998. Since 2000, he has been Vice Chairman of the IEEE Robotics and Automation Taipei Chapter. He won the Outstanding Research Award granted by National Science Council (NSC), Taiwan, in 1997 and 1999, and the Outstanding Electrical Engineering Professor Award granted by the Chinese Institute of Electrical Engineering (CIEE) in 1997, and the Outstanding Engineering Professor Award granted by the Chinese Institute of Engineering in 2000. He was also elected as one of the 38th Ten Outstanding Young Persons in Taiwan, R.O.C., in 2000.