THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 54, No. 9, November 1975
Printed in U.S.A.

A Comparative Study of Various Quantization Schemes for Speech Encoding

By P. NOLL

(Manuscript received February 7, 1975)

In this paper, the performance limits, as given by the signal-to-noise ratio (s/n), are described for different speech-encoding schemes including adaptive quantization and (linear) adaptive prediction schemes. The comparison is made on the basis of computer simulations using 8-kHz-sampled speech signals of one speaker. Different bit rates (two bits per sample—five bits per sample) have been used.

A three-bit-per-sample PCM scheme with a nonadaptive μ100 quantizer leads to an s/n value of approximately 9 dB. A maximum s/n value of approximately 25 dB has been reached using an encoding scheme including both adaptive quantization and adaptive prediction. Entropy coding of the quantizer output symbols leads to an additional gain in s/n of nearly 3 dB.

I. INTRODUCTION

Design of an efficient encoding scheme requires some knowledge of the statistics of the signal. Efforts to improve the performance of PCM systems have taken two primary directions:

(*i*) Use of quantizing schemes based on knowledge of the (one-dimensional) probability density function (PDF) of the samples to be quantized.
(*ii*) Use of quantizing schemes exploiting the correlation between successive samples.

If we had an *a priori* knowledge of the statistics of the samples, a nearly optimum quantization scheme could be used consisting of:

(*i*) A quantizer matched to the PDF of the signal to be quantized.
(*ii*) A predictor optimized for the given autocorrelation function of the signal.

The predictor lowers the variance of the signal to be quantized by

removing the correlation between successive samples. This is done by subtracting an estimation value from each incoming sample; the difference can be quantized, encoded, and transmitted (differential PCM = DPCM).

In digital speech-encoding systems, we have only a small amount of *a priori* knowledge of the statistics which, in addition, usually change with time:

(*i*) The long-period mean level differs from speaker to speaker.
(*ii*) At a given mean level, the instantaneous level changes because of variations in speech sounds.
(*iii*) The correlations between successive samples change because of variations in speech sounds.

To overcome these problems of unknown statistics, adaptive quantization and adaptive prediction schemes must be used. In these schemes, local estimates of the statistical parameters are calculated. The quantizer and/or predictor are then optimized based on these estimates.[1–3]

This paper compares different encoding schemes that include:

(*i*) Fixed quantizers.
(*ii*) Adaptive quantizers.
(*iii*) Fixed predictors.
(*iv*) Adaptive predictors.

The comparison is done on the basis of computer simulations; the signal-to-quantization noise ratio (s/n) has been used as the criterion for the comparisons. It is believed, however, that the s/n understates the subjectively perceived performance of encoders that have differential quantizers (the DPCM schemes in this paper).[1]

II. DESCRIPTION OF THE ENCODING SCHEMES

A computer program has been written that allows the simulation of encoding schemes combining the possibilities of nonadaptive or adaptive quantization and nonadaptive or adaptive prediction. The schemes that have been used are described in the following sections.

2.1 Fixed and adaptive quantizers

If the quantizer is *nonadaptive*, its characteristic is assumed to be logarithmic. Optimum, i.e., s/n-maximizing quantizers (whether uniform or nonuniform), cannot be used, not even under the assumption of a constant mean level, because the idle channel noise is higher for op-

timum quantizers than for logarithmic quantizers and results in poorer subjective performance.[3,4] The idle channel noise performance is determined by the smallest reconstruction level r_1 of the quantizer. Table I lists these values for various optimum three-bit quantizers (the term Gauss quantizer refers to a quantizer with an s/n-maximizing performance for signals with a gaussian PDF, etc.). The gamma PDF is a good model for speech amplitudes, but the smallest reconstruction level is 2.4 times higher for the corresponding optimum quantizer than for the logarithmic quantizer.

To overcome the problems of unknown mean level and the variations of the instantaneous level, adaptive quantization schemes (AQ schemes) can be used. A local estimate σ_x^2 of the variance of the input signal can be calculated; this value controls the gain of an amplifier located in front of a quantizer that is optimum for signals with unit variance. Two schemes are possible:

(*i*) *Forward estimation* (AQF): The estimation value is calculated from samples of the input signal. The input signal must be buffered, and the estimation value must be transmitted to the receiver in addition to the quantized samples.[3,5]

(*ii*) *Backward estimation* (AQB): The estimation value is calculated from quantized samples[3–6]; therefore, the state of the amplifier need not be transmitted (except for synchronizing purposes in case of channel errors).

Figure 1 shows the structures of the different PCM schemes. Note that the combination of controlled amplifier and fixed quantizer can be replaced by a quantizer with a step-size adaptation. Matching the gain of the amplifier to signal variance results in modifying the PDF of the signal to be quantized. It has been shown that different density functions can be reached by choosing an appropriate forward estimation scheme.[3] To get the best s/n performance, those quantizers can be employed that are optimum for the specific PDF.

Table I — Comparison of the smallest reconstruction levels r_1 of different optimum unit variance three-bit quantizers

Type of Quantizer	Nonuniform Quantizer r_1	Uniform Quantizer r_1
Uniform PDF quantizer	—	0.217
Gauss quantizer	0.245	0.293
Laplace quantizer	0.222	0.366
Gamma quantizer	0.149	0.398
Logarithmic μ100 quantizer	0.062	—

2.2 Types of quantizers

The following types of quantizers were used in the simulation of the speech-encoding systems:

Uniform quantizer with different loading factors.
Logarithmic quantizer with different loading factors.

Uniform optimum Gauss quantizer.
Uniform optimum Laplace quantizer.
Uniform optimum gamma quantizer.

Nonuniform optimum Gauss quantizer.
Nonuniform optimum Laplace quantizer.
Nonuniform optimum gamma quantizer.

These optimum quantizers lead to a maximum s/n for the specific probability density functions.

2.3 Algorithms of the AQ schemes

In applying adaptive quantization schemes, different possibilities of controlling the gain of the amplifier have been used. The following notation has been employed for the description of the algorithms (see also Figs. 1 and 2):

Symbol	Explanation
$x(n)$	Input sample at time instant n.
$y(n)$	Quantized sample at time instant n.
I_n	Index of quantizer step at time instant n.
G_n	Gain of the amplifier at time instant n (backward estimation).
G_N	Gain of the amplifier used in block N (forward estimation).
M	Number of quantized samples used for calculation of G_n (backward estimation).
N	Number of block.
$NSEG$	Number of input samples used for calculation of G_N (forward estimation).
NF	Number of first sample of block N. $NF = (N-1)\cdot NSEG + 1$.
ϱ_N	Vector of short-term autocorrelation coefficients calculated from the input samples of block N.
$\mathbf{R}_N$	Toeplitz matrix of short-term autocorrelation coefficients calculated from the input samples.
$\alpha, \beta, \alpha_j, \epsilon$	Coefficients to be optimized for each algorithm.

In all AQ schemes, a local estimate of the quantizer input signal variance

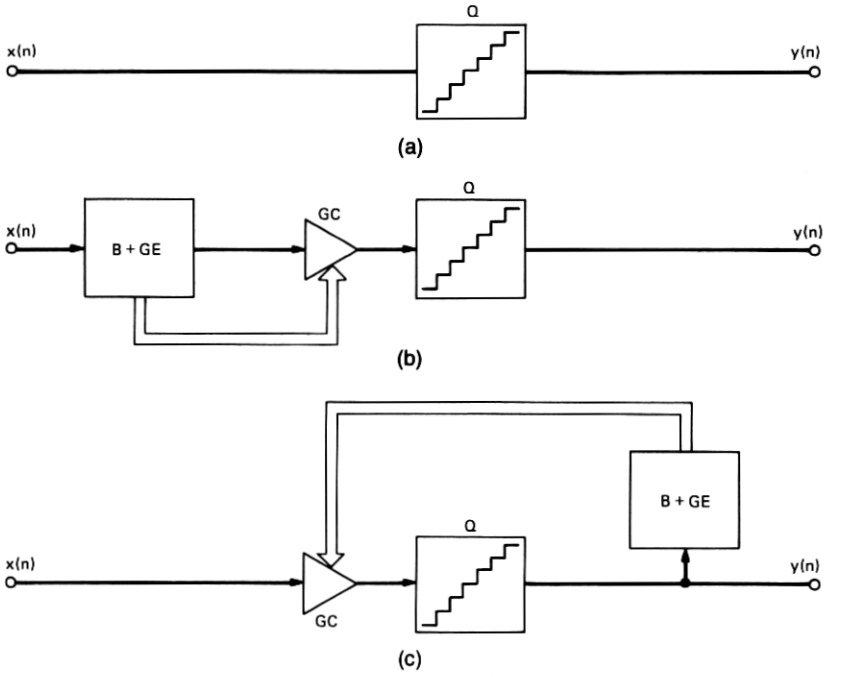

Fig. 1—PCM encoding schemes. (a) nonadaptive PCM. (b) Adaptive PCM with forward estimation (PCM-AQF). (c) Adaptive PCM with backward estimation (PCM-AQB). $x(n)$ = input signal, $y(n)$ = quantized signal, Q = quantizer, GC = gain control, B + GE = buffer and gain estimation.

is calculated; this value determines the gain of the amplifier such that the quantizer is optimal loaded.

2.3.1 Forward estimation schemes (AQF)

In the AQF schemes, the gain is only readjusted once for a new block of $NSEG$ speech samples:

$$G_N = \text{const.}; \qquad n = NF, NF + 1, \cdots, NF + NSEG - 1$$
$$NF = (N - 1) \cdot NSEG + 1.$$

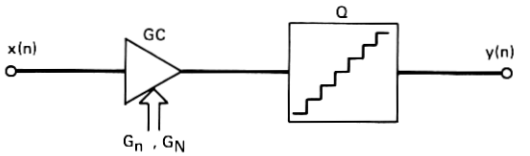

Fig. 2—AQ scheme. $x(n)$ = input signal, $y(n)$ = quantized signal, G_n = gain used at time instant n (AQB scheme), G_N = gain used in block N (AQF scheme).

The following algorithms have been used:

(*i*) PCM: *variance scheme*[3]: An unbiased estimation of the variance of the block is calculated:

$$G_N^{-2} = \alpha \cdot \frac{1}{NSEG} \sum_{j=1}^{NSEG} x^2(NF - 1 + j). \tag{1}$$

G_N is proportional to the inverse of the standard deviation estimated from the samples of the block.

(*ii*) PCM: *maximum scheme*: The maximum amplitude in the block is used:

$$G_N^{-1} = \alpha \cdot \max \{|x(NF - 1 + j)|\}_{j=1,\cdots,NSEG}. \tag{2}$$

(*iii*) DPCM: *maximum scheme*[7]: The maximum difference between neighbored samples is used:

$$G_N^{-1} = \alpha \cdot \max \{|x(NF - 1 + j) - x(NF - 2 + j)|\}_{j=2,\cdots,NSEG}. \tag{3}$$

This algorithm can only be used for predictors with one coefficient.

(*iv*) ADPCM: *variance scheme*[8,9]: The vector of short-term autocorrelation coefficients is used to calculate an estimation value of the variance σ_d^2 of the difference signal:

$$G_N^{-2} = \alpha \cdot \sigma_d^2 = \alpha \cdot [\sigma_x^2 - \varrho_N^T \cdot \mathbf{R}_N^{-1} \cdot \varrho_N] \tag{4}$$

2.3.2 Backward estimation schemes (AQB)

In AQB schemes, the gain of the amplifier is, in general, modified for every new input sample by a factor depending on the knowledge of the previous quantized samples or of the corresponding quantizer indices.

$$G_n = \alpha \cdot G_{n-1}.$$

The following algorithms have been used:

(*i*) *One-word memory scheme*[6]: The last gain value is multiplied by a factor that depends on the last occupied quantizer step:

$$G_n = f(|I_n|)G_{n-1}. \tag{5}$$

(*ii*) *Variance scheme*[6,10]: The last M quantized samples and (for $\beta \neq 0$) the last gain value are used to calculate a new gain value:

$$G_n^{-2} = \sum_{j=1}^{M} \alpha_j y^2(n - j) + \beta / G_{n-1}^2. \tag{6}$$

(*iii*) *Modified one-word memory scheme*[3]: The gain of the amplifier is changed if the smallest reconstruction level has been occupied α times or if the largest reconstruction level has been occupied once:

$$G_n = \begin{cases} 2.0 \cdot G_{n-1} & \text{if } |I_m| = \min \text{ for } \alpha \text{ times } (m = n, n-1, \cdots, n-\alpha+1) \\ 0.5 \cdot G_{n-1} & \text{if } |I_n| = \max \\ 1.0 \cdot G_{n-1} & \text{otherwise.} \end{cases} \tag{7}$$

2.4 Fixed and adaptive predictors

In predictive encoding systems, an estimate of each input sample is calculated and subtracted from the actual input sample; the difference is then quantized, encoded, and transmitted. The use of nonadaptive predictors (DPCM schemes) leads to a suboptimum overall performance of the encoding scheme, because the prediction is not optimum for all speakers and for all speech sounds.[1,8]

A better prediction can be reached by using adaptive algorithms (ADPCM schemes). Two schemes are possible:

(*i*) *Forward scheme*: A short-time autocorrelation function is calculated using a finite number of buffered input samples. The predictor coefficients are readjusted according to the time-variant autocorrelation function.[2,11]

(*ii*) *Backward scheme*: The predictor is optimized using the quantized information (gradient search method and Kalman filter algorithm).[1,2]

Only the forward scheme has been used in the simulations. The optimum vector $\mathbf{h}_N$ of J predictor coefficients for each block N is

$$\mathbf{h}_N = \mathbf{R}_N^{-1} \cdot \boldsymbol{\varrho}_N. \tag{8}$$

$\mathbf{R}_N$ and $\boldsymbol{\varrho}_N$ are the matrix and the vector of short-term autocorrelation coefficients calculated from the input samples of block N. The predictor coefficients have to be transmitted to the receiver, in addition to the code words of the quantized difference signal samples. An upper bound of the gain in s/n as compared to PCM is given in Section IV.

III. RESULTS

Various nonadaptive and adaptive encoding schemes have been simulated on a digital computer. The signal-to-quantization noise ratio (s/n) has been determined using 8-kHz-sampled speech samples of one speaker. The same 2.3-s utterance ("The boy was mute about his task"; female voice; bandwidth 200 to 3200 Hz) has been used in

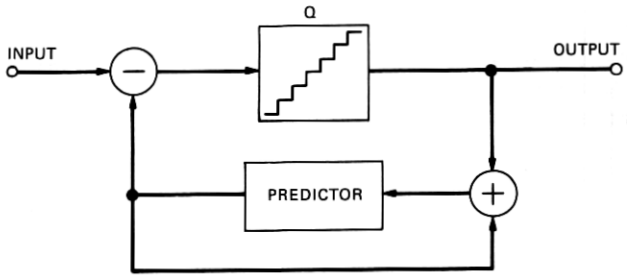

Fig. 3—DPCM scheme. Q = quantizer.

all simulations. (The simulations have not included any high-frequency emphasis of the input speech, as is characteristic of a 500-type set transmitter, for example.) The following schemes have been studied:

(*i*) Nonadaptive Quantization
PCM: see Fig. 1.
DPCM (nonadaptive prediction): see Fig. 3.
ADPCM (adaptive prediction): see Fig. 4.

(*ii*) Adaptive Quantization
PCM: see Fig. 1.
Forward scheme (PCM-AQF).
Backward scheme (PCM-AQB).
DPCM (nonadaptive prediction): see Fig. 5.
Forward scheme (DPCM-AQF).
Backward scheme (DPCM-AQB).
ADPCM (adaptive prediction): see Fig. 6.
Forward scheme (ADPCM-AQF).
Backward scheme (ADPCM-AQB).

These encoding schemes have been optimized using the s/n as criterion.

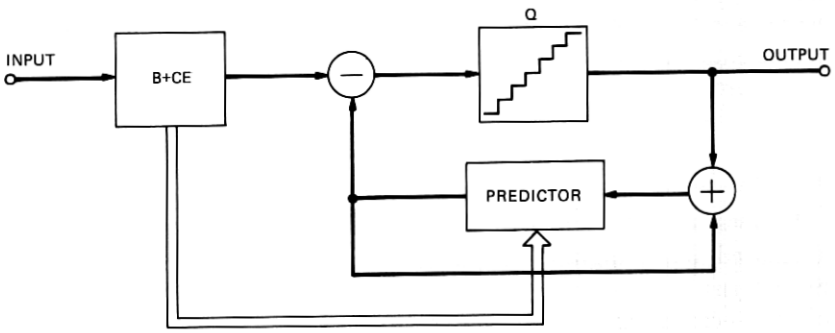

Fig. 4—ADPCM scheme. B + CE = buffer and coefficients estimator.

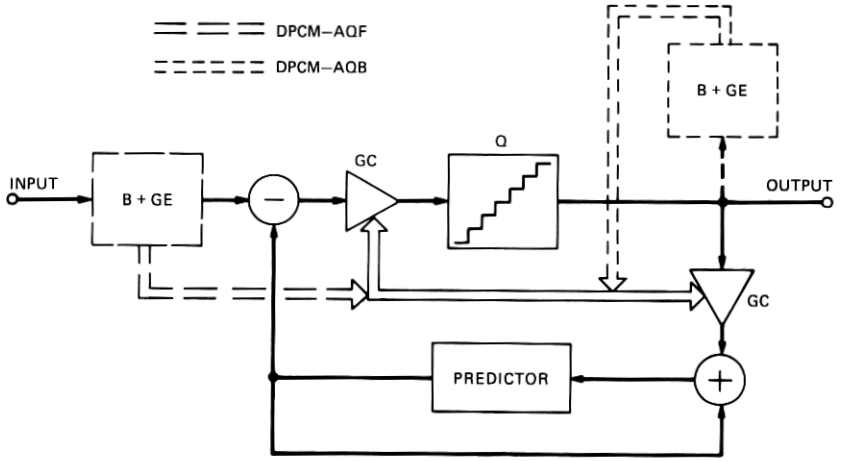

Fig. 5—DPCM-AQ schemes. B + GE = buffer and gain estimator, GC = gain control.

3.1 Optimum results: three bits/sample quantization

Figure 7 shows the optimum results reached with a three-bit quantization of the 2.3-s speech sample.

Left curves: Optimum results using a fixed quantizer.
Right curves: Optimum results using an adaptive quantizer.
Lower curves: Prediction with a first-order predictor (one coefficient).
Upper curves: Prediction with a high-order predictor.

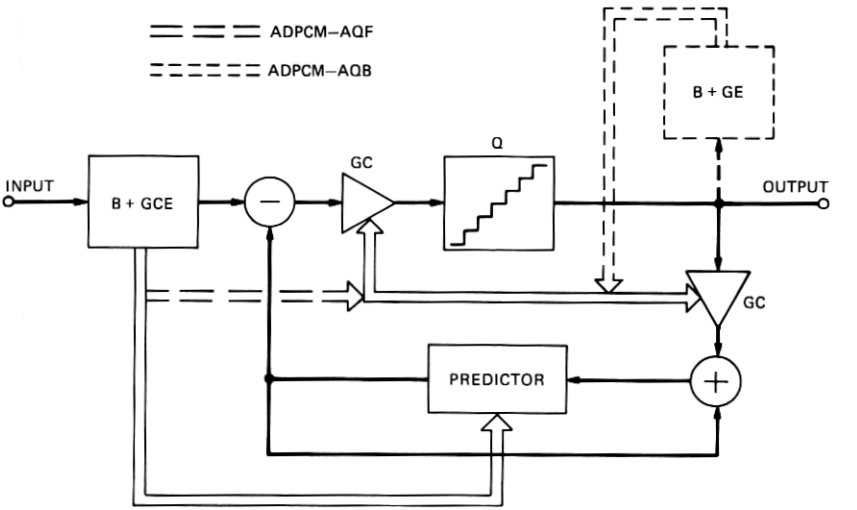

Fig. 6—ADPCM-AQ schemes. B + GCE = buffer and gain and coefficients estimator.

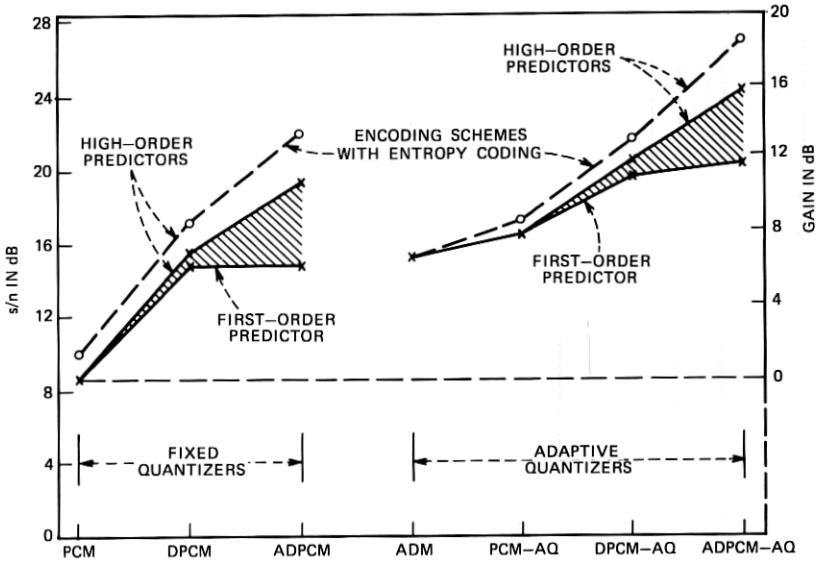

Fig. 7—Signal-to-noise ratio values and gains (over logarithmic PCM) for different three-bit speech-encoding systems.

Quantizers with a logarithmic characteristic have been used in all simulations with a fixed quantizer (curves on the left side of Fig. 7). The s/n value for a PCM scheme is

$$\text{s/n} = 8.7 \text{ dB}$$

if the quantizer has a $\mu 100$ characteristic, and if the loading is $4\sigma_x$ (σ_x is the standard deviation of the signal to be quantized). As compared to this s/n value of 8.7 dB, the following maximum gains can be reached with prediction schemes using the same type of quantizer (G^* is the gain in s/n over PCM):

$$\text{Fixed predictor, fixed quantizer: } G^* \approx 7 \text{ dB}$$
$$\text{Adaptive predictor, fixed quantizer: } G^* \approx 11 \text{ dB}.$$

Adaptive quantization (PCM-AQ) not only has the advantage of increasing the dynamic range that the quantizer can handle, but it also allows the application of quantizers that are optimum for the probability density function of the signal to be quantized. The following gain over the 8.7-dB value of nonadaptive PCM has been reached:

$$\text{Adaptive quantization: } G^* \approx 7 \text{ dB}.$$

Using predictors, the gains over PCM are now

$$\text{Fixed predictor, adaptive quantizer: } G^* \approx 12 \text{ dB}$$
$$\text{Adaptive predictor, adaptive quantizer: } G^* \approx 16 \text{ dB}.$$

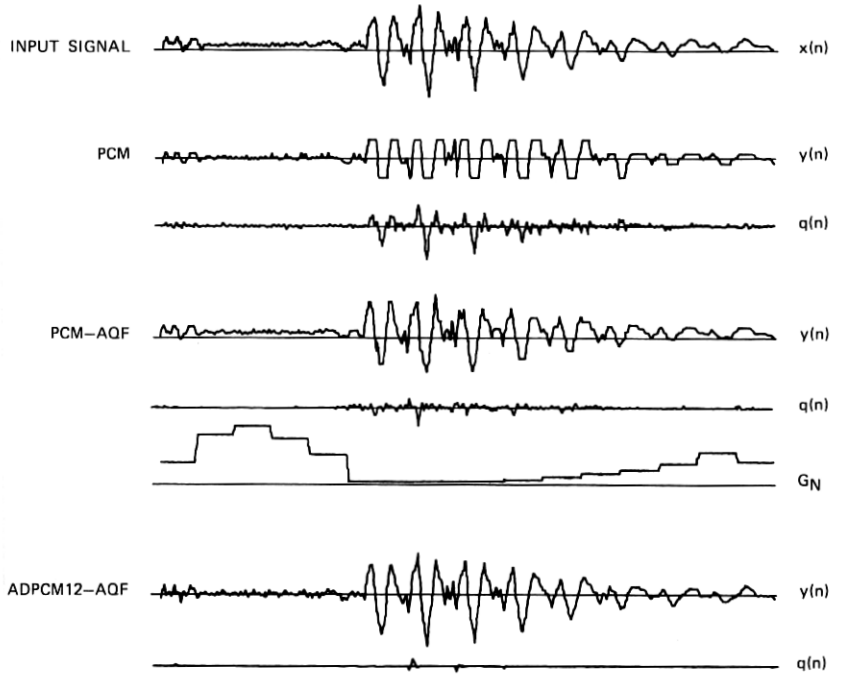

Fig. 8—Comparison of waveforms. $x(n)$ = sequence of input samples (512 samples), $y(n)$ = sequence of decoded samples, $q(n)$ = sequence of quantization errors, G_N = sequence of amplifier gains.

Figure 8 shows the waveforms of the reconstructed signal and of the quantization error for a 64-ms segment of speech. Three examples are shown:

(*i*) PCM, nonadaptive, $\mu 100$ characteristic. Only eight different levels can be used for the reconstruction (decoding) of the signal.

(*ii*) PCM-AQF, optimum Gauss quantizer, $NSEG = 32$. The number of levels is limited to eight for each segment of $NSEG$ samples. Different levels can be used for each segment.

(*iii*) ADPCM12-AQF, optimum Laplace quantizer, $NSEG = 128$. The predictive encoding with a 12th-order predictor leads to a very high s/n. For each segment, the number of levels of the *difference* signal is limited to eight, but the reconstructed signal does not suffer this limitation.

3.2 Adaptive delta modulation

To determine whether the quantization schemes represent an improvement over existing adaptive delta modulation (ADM) schemes, the s/n value of Jayant's ADM-scheme[12] has been determined at a bit rate

of 24 kb/s. The s/n value is approximately 15 dB.[7] Therefore, the gain over nonadaptive three bits/sample PCM is

$$\text{Adaptive delta modulation: } G^* \approx 6 \text{ dB}.$$

3.3 Entropy coding

Entropy coding is a variable-length coding procedure that assigns short code words to highly probable symbols and longer code words to less probable symbols. The average word length is approximately equal to the entropy of the quantizer output signal. The entropy coding technique leads to an additional gain in s/n for a given average bit rate. The number of quantizer steps can be increased without exceeding an average bit rate of three bits per sample. The dashed lines in Fig. 7 show the s/n values that can be reached by using an entropy coding technique. In this case, uniform quantizers with a large number of steps have been employed; the step sizes have been adjusted to give a quantizer output entropy of three bits. It should be noted that a buffer is needed so that the variable-length coded signal can be transmitted over a channel at a uniform bit rate.

3.4 Optimum results: two bits/sample up to five bits/sample quantization

Figure 9 shows the s/n values for quantizations with two bits/sample up to five bits/sample (corresponding to bit rates from 16 kb/s up to 40 kb/s). The following encoding schemes have been compared:

PCM	μ100 characteristic, $8\sigma_x$ loading.
PCM-AQF	$NSEG = 32$, optimum Gauss quantizer.
DPCM1-AQB	1 predictor-coefficient, fixed; optimum Gauss quantizer.
ADPCM1-AQF	1 predictor-coefficient, adaptive; optimum Gauss quantizer; $NSEG = 32$.
ADPCM4-AQF	4 predictor-coefficients, adaptive; optimum Laplace quantizer; $NSEG = 128$.
ADPCM12-AQF	12 predictor-coefficients, adaptive; optimum gamma quantizer; $NSEG = 256$.

3.5 Parameter transmission in adaptive encoding schemes

In all forward schemes, channel capacity is needed for transmission of the adaptive parameters. The problems and techniques of quantizing these parameters are not considered in this paper. It is known that the parameters tolerate coarse quantization and slow updating. If necessary, redundancy-reducing schemes can be used to lower the number of bits that have to be transmitted in addition to the encoded speech

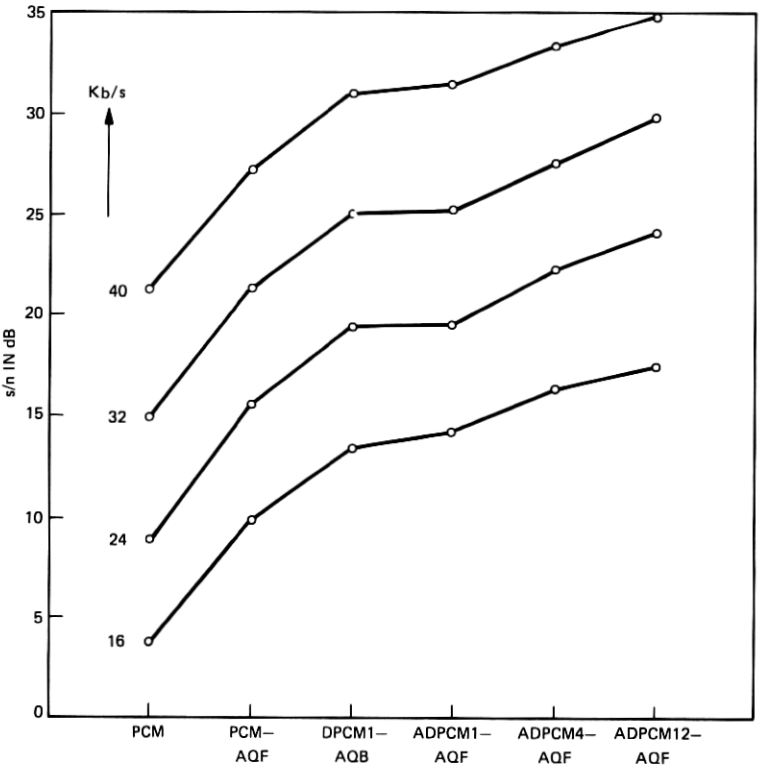

Fig. 9—Signal-to-noise ratio values for quantization with two bits per sample (16 kb/s) up to five bits per sample (40 kb/s).

samples. The needed channel capacity can be approximately transformed into an equivalent loss in s/n. If each parameter of the adaptive scheme has to be encoded with $NADD$ bits/segment, and if $NSEG$ is the number of samples/segment, then we get an equivalent reduction in s/n performance:

$$\Delta_{s/n} = 6.02 \frac{NADD \text{ (bits/segment)}}{NSEG \text{ (samples/segment)}} \text{ (dB)}. \tag{9}$$

This loss is due to the reduction of the number of quantizer steps in order not to exceed the maximum allowed bit rate.

Example:

$NSEG = 128$ (16 ms)
$NADD = 4$ bit.

The loss is 0.2 dB for each coefficient to be transmitted.

IV. UPPER BOUNDS FOR PREDICTION

The linear dependencies between the amplitudes of the speech sample being used in all simulations have been calculated to get a measure of the maximum gain that can be reached with linear prediction. Note that these upper bounds of the prediction gain cannot be reached with predictive encoding systems (especially if the quantizer has only a low number of quantization levels), because prediction is done then with decoded speech samples. These samples include a quantization error.

4.1 Nonadaptive prediction

The long-term autocorrelation function of the speech signal has been measured. Figure 10 shows the first 19 time lags of the normalized autocorrelation function $\rho(n)$. Using these data, a predictor can be optimized such that the variance of the difference signal is minimum.

The prediction gain is the ratio of the variances of the input signal and the difference signal:

$$G_P = 10 \log_{10} \frac{E[x^2(n)]}{E[d^2(n)]} = 10 \log_{10} \frac{\sigma_x^2}{\sigma_d^2}. \tag{10}$$

G_P can be calculated directly from the normalized autocorrelation function $\rho(n)$. Figure 11 shows this gain versus the number of coefficients being used for the prediction. The maximum prediction gain is approximately 10.5 dB. This value is an upper bound of the additional gain in s/n over PCM by using nonadaptive differential encoding schemes. This gain cannot be reached if the DPCM encoder has to handle speech samples of different speakers. In this case, suboptimum predictor coefficients have to be chosen such that the DPCM encoder has

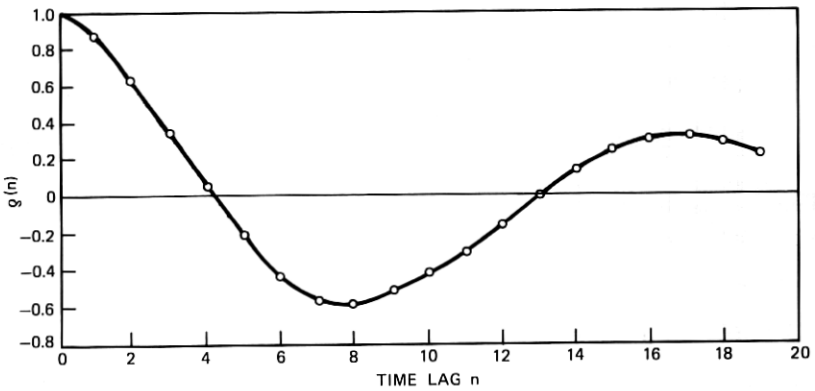

Fig. 10—Normalized autocorrelation function (female voice; 200 to 3200 Hz).

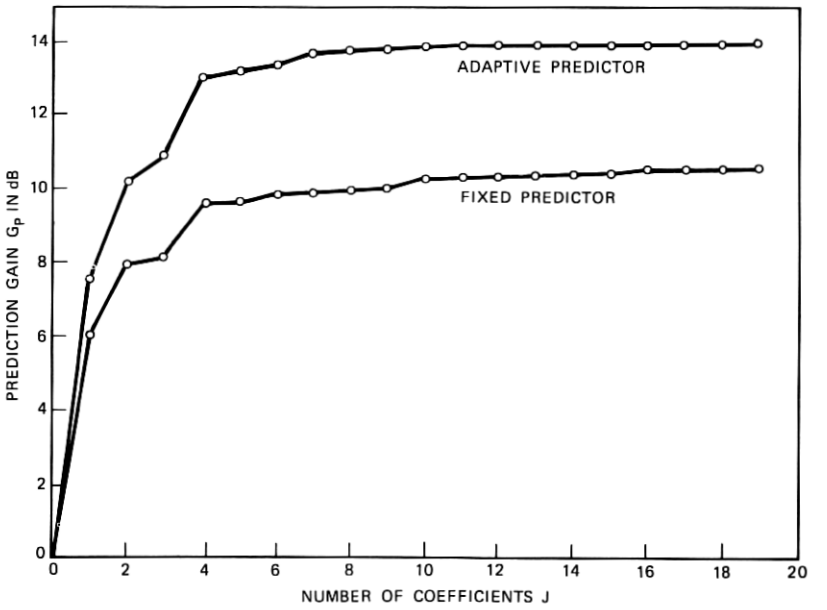

Fig. 11—Prediction gains vs number of predictor coefficients.

a good performance for all speakers. This demand can only be fulfilled with predictors of low order (up to three coefficients). It may be relevant to mention that such suboptimum predictor coefficients have been used in the simulation of the DPCM schemes.

Knowing the long-term autocorrelation function $\rho(n)$, it is possible to calculate an approximation of the power density function. This is done by calculating the power transfer function of a recursive filter, the coefficients of which are equivalent to the coefficients of the optimum predictor (maximum-entropy method[13]). Figure 12 shows the power density spectrum calculated in this way from 16 coefficients of the autocorrelation function $\rho(n)$.

4.2 Adaptive prediction

NSEG samples of the input samples are buffered, and the short-term autocorrelation function of this segment is calculated. For each segment of *NSEG* samples, the variance of the difference signal can be calculated directly from this short-term autocorrelation function [see eq. (4)]. Using these variances, a prediction gain can be determined for different numbers of predictor coefficients and for different values of *NSEG*. Figure 11 shows the optimum prediction gain for an adaptive prediction scheme versus the number of predictor coefficients. In each

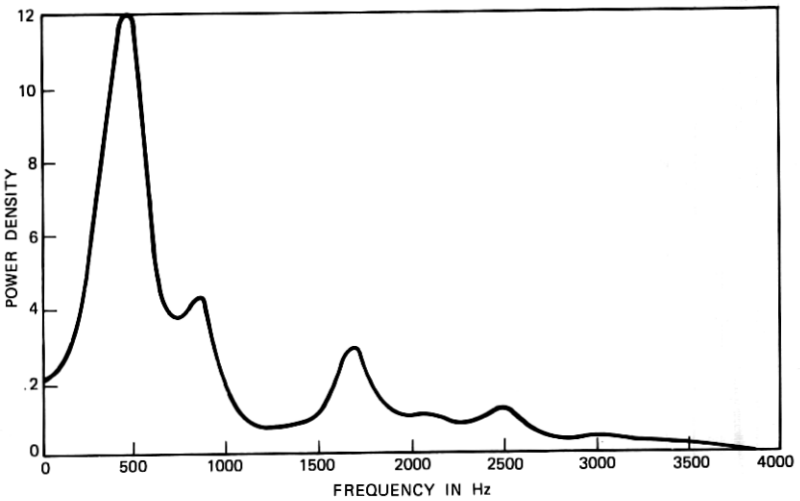

Fig. 12—Power density of the speech signal.

case, the optimum value *NSEG* has been used. The maximum prediction gain is approximately 14 dB. This value is an upper bound of the additional gain in s/n over PCM by using adaptive differential encoding schemes.

V. UPPER BOUNDS FOR QUANTIZATION

It is possible to design quantizers such that the signal-to-quantizing-noise ratio is a maximum; this is done by choosing the quantizer step sizes according to the probability density function of the signal. It is known that these optimum quantizers cannot be used for the quantization of speech signals: the s/n improvement is offset by the greater idle channel noise and smaller dynamic range (Ref. 4; see also Section 2.1 above). Optimum quantization is practical, however, if used in an adaptive quantization scheme; it gives us an s/n advantage over logarithmic quantization, and it allows a further increase in s/n by using entropy coding techniques (variable length coding). The adaptive quantization technique changes the PDF of the signal to be quantized; it has been shown[3] that different density functions can be reached with the forward estimation scheme (AQF scheme). Table II shows the s/n values for three-bit quantizers without and with entropy coding. The values of the first two columns are taken from Max[14] and Paez and Glisson.[15] In the case of entropy coding, the quantizers have been optimized so that the s/n is maximum for the given average bit rate of three bits per sample.[16] It is not possible to get higher s/n values with any encoding scheme based on memoryless single-letter quantization.

Table II — Maximum s/n values of various three-bit quantizers

	Quantizer Without Entropy Coding		Quantizer With Entropy Coding
	Optimum Uniform Quantizer	Optimum Nonuniform Quantizer	
	s/n(dB)	s/n(dB)	s/n(dB)
Uniform PDF	18.06	(18.06)	18.06
Gaussian PDF	14.27	14.62	16.53
Laplace PDF	11.44	12.61	17.09
Gamma PDF	8.78	11.47	18.78

VI. CONCLUSIONS

Comparisons of various nonadaptive and adaptive three-bit speech-encoding systems via simulation with speech inputs show that a wide range of signal-to-quantization-noise ratios can be reached starting with 9 dB (logarithmic PCM) and increasing up to 27 dB (adaptive predictive coding with adaptive quantization and entropy coding). Adaptive quantization has an s/n advantage of 7 or 5 db over logarithmic PCM when used in encoding schemes without and with prediction, respectively. Nonadaptive prediction leads to a 7-dB increase in s/n, and 11 dB can be gained using adaptive prediction techniques. Entropy coding gives an additional 2 to 3 dB improvement; such a coding technique is difficult to implement if a constant bit rate has to be achieved, but it may be of interest for asynchronous data networks. Furthermore, subjectively, DPCM gains over logarithmic PCM are believed to be greater than what the s/n gains suggest.[1]

Informal listening tests have shown that all predictive encoding schemes give a very good speech quality when used in connection with adaptive quantization (DPCM-AQ or ADPCM-AQ). Differences between the original speech and the decoded speech are not audible with adaptive prediction schemes when a high-order predictor is used (for example, ADPCM4-AQF).

The upper bounds that have been determined separately for the prediction gains and the quantizer s/n performances cannot be reached in practical predictive encoding systems. This fact is attributed to the predictor-quantizer interaction; that is, the input to the predictor is a noisy version of the input signal, and the input to the quantizer is a noisy prediction error. This interaction is not negligible when three-bit quantizers are used.

It is important to realize that all results are based on a single speech record of one speaker. Computer simulations using other speech

material show basically similar results; the main differences appear in the exact prediction gains that can be reached. In many instances, these gains are higher than those mentioned in this paper.

One object of this paper was to quantify the (relative and absolute) capabilities of a wide range of nonpitch-tracking speech coders. The coders studied have a variety of potential applications that call for different specifications of speech quality and coder complexity. A second purpose of this paper was to study the capabilities of three-bit encoding in some detail, as motivated by mobile telephone studies.[17,18]

REFERENCES

1. N. S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers," Proc. IEEE, *62*, No. 5 (May 1974), pp. 611–632.
2. P. Noll, "Untersuchungen zur Sprachcodierung mit adaptiven Prädiktionsverfahren," Nachrichtentechnische Zeitschrift, *27*, H. 2 (1974), pp. 67–72.
3. P. Noll, "Adaptive Quantizing in Speech Coding Systems," Proc. of the International Zürich Seminar on Digital Communications, 1974, pp. B3(1)–(6).
4. R. W. Stroh and M. D. Paez, "A Comparison of Optimum and Logarithmic Quantization for Speech PCM and DPCM Systems," IEEE Trans. on Commun., *COM-21*, No. 6 (June 1973), pp. 752–757.
5. R. Elsner, W. K. Endres, H. Mangold, P. Noll, E. Paulus, and D. Wolf, "Recent Progress in Digital Processing of Speech," IEEE Trans. on Commun., *COM-22*, No. 9 (1974), pp. 1168–1171.
6. P. Cummiskey, N. S. Jayant, and J. L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," B.S.T.J., *52*, No. 7 (September 1973), pp. 1105–1118.
7. N. S. Jayant, private communication.
8. P. Noll, "Nonadaptive and Adaptive Differential Pulse Code Modulation of Speech Signals," Polytech. Tijdschr., Ed. Elektrotech. Elektron. No. 19 (1972), pp. 623–629.
9. P. Noll, "Sprachübertragung mit adaptiven DPCM-Verfahren," Heinrich-Hertz-Institut, Berlin, Germany, Tech. Bericht 164, 1973.
10. R. W. Stroh, "Optimum and Adaptive Differential Pulse Code Modulation," Ph.D. thesis, Polytechnic Institute of Brooklyn, 1970.
11. B. S. Atal and M. R. Schroeder, "Adaptive Predictive Coding of Speech Signals," B.S.T.J., *49*, No. 8 (October 1970), pp. 1973–1986.
12. N. S. Jayant, "Adaptive Delta Modulation with a One-Bit Memory," B.S.T.J., *49*, No. 3 (March 1970), pp. 321–343.
13. J. P. Burg, "Maximum Entropy Spectral Analysis," Proc. NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics, Enschede (Netherlands), 1968.
14. J. Max, "Quantizing for Minimum Distortion," IRE Trans. Inform. Theory, *IT-6*, No. 1 (March 1960), pp. 7–12.
15. M. D. Paez and T. H. Glisson, "Minimum Mean-Squared-Error Quantization in Speech PCM and DPCM Systems," IEEE Trans. on Commun. Tech., *COM-20*, No. 2 (April 1972), pp. 224–230.
16. P. Noll and R. Zelinski, "Ein Beitrag zur Quantisierung gedächtnisfreier Modellquellen," Heinrich-Hertz-Institut Berlin-Charlottenburg, Tech. Report No. 170, 1974.
17. P. Noll, "Effects of Channel Errors on the Signal-to-Noise Performance of Speech-Encoding Systems," B.S.T.J., this issue, pp. 1615–1636.
18. N. S. Jayant, "Step-Size Transmitting Differential Coders for Mobile Telephony," B.S.T.J., this issue, pp. 1557–1581.