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

Step-Size Transmitting Differential Coders for Mobile Telephony

By N. S. JAYANT

(Manuscript received February 7, 1975)

In using digital speech for mobile radio, we encounter the problem of severe bit-error bursts. Error clustering occurs because the bit duration is typically much smaller than that of a signal "fade," and average bit-error probabilities greater than 1 percent are not uncommon. For speech communication over such channels, this paper proposes variable step-size differential coders based on explicit (and error-protected) transmission of quantizer step size. Specifically, we discuss delta and DPCM coders to be referred to as DM-AQF and DPCM-AQF, where AQF stands for adaptive quantization with forward estimation (and transmission) of step size. (Backward estimation, based on quantized-signal history, has the nice feature that the step-size information does not have to be explicitly transmitted. Furthermore, obtaining this information does not entail any encoding delay. However, due to the dependence of step size on reconstructed signal history, backward estimation is often less reliable in the presence of bit errors than a scheme based on AQF.) The studies reported in this paper cover the problem of step-size determination in AQF, the design of time-invariant first-order predictors for DPCM-AQF, and the performances of AQF encoders with and without burst-error-protecting ploys such as redundant time-diversity coding and bit scrambling. Judging from SNR figures and informal listening tests, interesting results are obtained with the following 48-kb/s coders: three-bit DPCM-AQF with redundant error protection, and DM-AQF using bit scrambling.

I. INTRODUCTION

Recent developments in speech digitization[1] have prompted an examination of digital coding as a possibility for mobile radio telephony that conventionally employs analog techniques for speech transmission. Conceivably, much of the signaling supervision and "bookkeeping" in a mobile radio link can be digital; in this case, if the speech were handled digitally as well, it would be simple to interleave the voice bits with the control bits for transmission. Digital coding also offers the possibilities of inexpensive coder-decoder implementation, straightforward speech encryption (by bit scrambling), and efficient signal regeneration. Perhaps the greatest incentive for the use of digital speech, however, is the thought that a properly designed digital code may be more resistant than analog systems to the multipath fading that characterizes mobile radio.

Figure 1 shows the envelope of a Rayleigh-fading signal that is typical in mobile telephony.[2] An important parameter is the fading rate, which is approximately the ratio of vehicle speed V to the carrier wavelength λ. For the example in Fig. 1, this ratio is about 15 Hz. Note also that the 5 m represent a total travel time of about 1 s at the indicated vehicle speed, and that the fading is slow or correlated in the sense that a given fade (signal strength below a specified threshold) can last for several tens of milliseconds (which will represent several hundred speech bits for the codes of this paper). The probability of a fade can be decreased by an order of magnitude by the use of diversity reception (two-branch, equal-gain or switched diversity, for example). But when a fade does occur, the signal is susceptible to noise capture as well as to co-channel interference. The end effects, with conventional analog transmissions, are impulsive "pops" and "crackles" in the

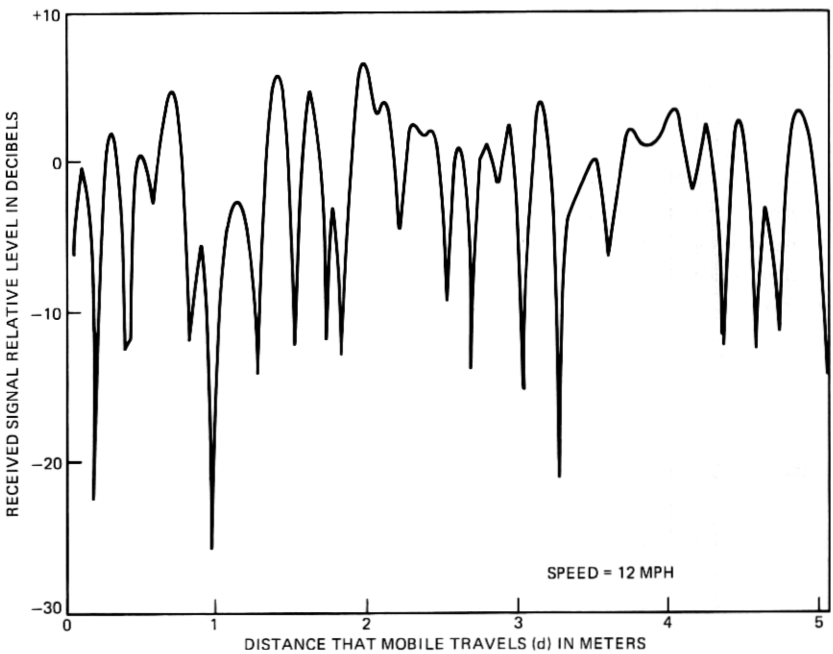

Fig. 1—Envelope of a Rayleigh fading signal ($V/\lambda = 15.4$ Hz).

received speech. Without explicit signal companding, these effects can be severe under "idle-channel" conditions, when no speech is present on the channel of interest. By contrast, if "adaptive" digital modulators are used to transmit speech, it is expected that the variable step-size mechanism in these modulators would inherently attenuate the impulsive interference signals (manifested as clustered bit errors in the digital system) during the silences of the incoming speech. This interference-squelching property is, in fact, expected to carry over to the "active-channel" condition with an ideally designed adaptive code (that exploits the statistics of signal fading or, equivalently, the time statistics of the bit-error bursts). Even if such an ideal design should be impractical, it is clear that digital coding can straightforwardly employ efficient burst-error-protection ploys such as time-diversity coding and bit-scrambling. However, these refinements (as well as the notion of forward estimation of step size) can involve significant amounts of encoding delay.

In our search for a suitable digital coder, we have used the following characteristics as guidelines. The speech bandwidth should be representative of standard telephone quality (200 to 3200 Hz); average bit-error rates higher than 1 percent are possible at times; and, finally, the overall transmission rate should not exceed a nominal 48 kb/s. When we refer to a "nominal 48-kb/s rate," we mean that additional channel capacity (in the order of 2 to 5 kb/s) may be needed for the transmission of step-size information.

A basic contention of this paper is that the "optimum" step size for a speech quantizer changes slowly enough with time for the step-size information to be transmitted reliably in a special error-protected format over a typical mobile radio channel. Thus, although the main stream of speech-carrying bits is still subject to errors, the provision of a relatively error-free step size will improve the received speech quality to a point that makes explicit step-size transmission worthwhile. We show that step-size transmitting coders are of interest for bursty as well as independent error patterns, and we include a comparison with a popular error-resistant syllabic-companded quantizer that recovers step size from the bit stream. Following Noll,[3] step-size transmitting adaptive coders will be labelled AQF (adaptive quantization with forward estimation and transmission of step size), in contradistinction to AQB (adaptive quantization with backward estimation).

The coders of this paper are differential. We discuss both DPCM (differential PCM) and DM (delta modulation). It appears from experience[1] that conventional time-invariant log-PCM quantization does not meet the error-performance requirements of mobile telephony. However, the possibility of a well-designed adaptive PCM[1,3,4] definitely exists. A promising candidate is the technique of nearly instantaneous companding (NIC).[4]

Although our studies have included informal perceptual assessments, most performance results in this paper are objective signal-to-noise ratios termed SNRT and SNRR. These reflect, respectively, the speech quality at the output of the local and remote decoders (T and R stand for "transmitter" and "receiver"). Formal definitions appear in Fig. 2. As we shall note at appropriate points in the paper, an SNRT-maximizing encoder does not, in general, maximize SNRR, and vice versa.

Our discussions refer to computer simulations that employed band-limited (200 to 3200 Hz) speech utterances (2 s or, sometimes, longer in duration) and bit-error patterns obtained from fading simulators.[5] We believe that the main conclusions of this paper should hold for broad classes of speech and error patterns encountered in a mobile radio environment. However, our numerical results are often reflective of the specific data used in our computer simulation. To demonstrate real-world variabilities of these numerical results, we have employed variable speech data, whenever appropriate.

Section II of this paper illustrates the time characteristics of the simulated burst error channel. Section III discusses the design of a DM-AQF coder. The section also demonstrates that simple bit-protecting codes are not particularly beneficial with DM-AQF (except for the transmission of step-size information). Bit scrambling, on the other hand, provides a definite advantage. Suitable sampling rates for DM-AQF are shown to be in the order of 30 to 40 kHz. Finally, a performance comparison is made between DM-AQF and a representative DM-AQB code. Section IV describes the design of a DPCM-AQF coder, and demonstrates

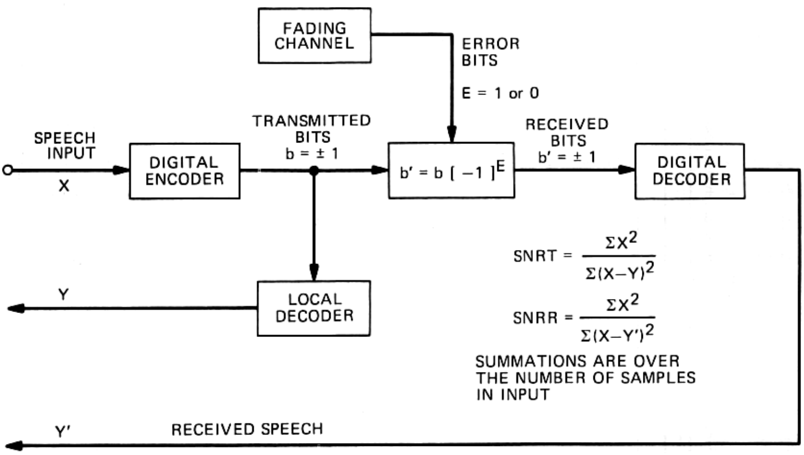

Fig. 2—Block diagram of codec simulation.

the utility of time-diversity coding for bit protection. It also indicates that a three-bit coder operating redundantly at a nominal 48 kb/s (and an 8-kHz sampling rate) is a better choice that a four-bit coder operating (also redundantly, but with less bit protection) at the same information rate. Section V provides a comparison of DPCM-AQF and DM-AQF.

II. BIT-ERROR PATTERNS

2.1 Burst errors

Two simulated-error sequences were used in this study, representing average error probabilities of 0.025 and 0.055. These numbers represent channel qualities believed to be typically "much worse than average."[5] The durations of the error sequences were long enough to simulate the transmission of all but the longest of the speech utterances being encoded. For this utterance (which was 9 s long), the bit-error sequences were used repeatedly to cover the total speech duration. Simulated bit rates ranged from 24 to 48 kb/s.

Figure 3 displays typical distribution functions for error-burst duration D and the error-free interval I. The numbers refer to a subsegment of the 0.025 error rate sequence. The error rate is denoted by $P(EB)$,

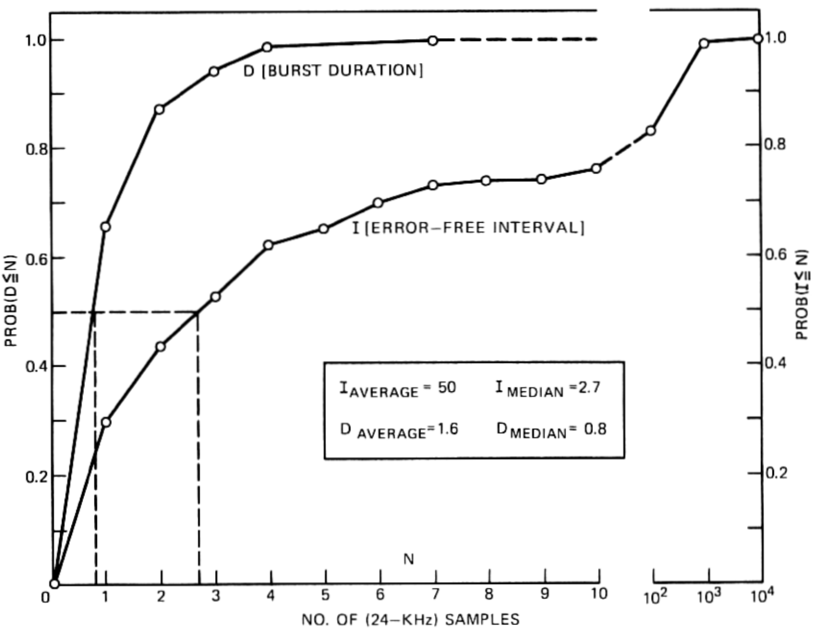

Fig. 3—Time statistics of error bursts [$P(E) = 0.025$, $V/\lambda = 36.2$ Hz].

where B refers to burst errors. An error burst in Fig. 3 is defined to have a local error probability of 1. In other words, an error burst of length D_0 implies that D_0 contiguous signal bits are in error. An isolated error, for example, is an error burst of length $D_0 = 1$. The upper rows of Table I lists the average and median values of burst duration D and error-free interval I for the subsequence examined. The ratio of (D_{average}) to $(D_{\text{average}} + I_{\text{average}})$ gives the average bit-error probability for the subsequence $(1.6/51.6 = 0.03)$. The ratio of (D_{median}) to $(D_{\text{median}} + I_{\text{median}})$, by contrast, is as high as 0.23. Of particular interest is the fact that $I_{\text{average}} \gg I_{\text{median}}$. This signifies the presence of some error-free intervals that are extremely long, together with a preponderance of intervals that are (unfortunately) very short (in fact, not much longer than D_{average}). The clustered nature of the errors is somewhat more apparent by comparison with average and median statistics that apply to an appropriate random error channel: that is, a channel where errors occur independently at every sample, but with an average error probability that is the same as that of the burst-error channel. The lower rows of Table I shows those statistics, as calculated for a random error channel whose bit-error probability is 0.03. Note that the value of D_{average} is much higher (for the same average error probability) in the case of the bursty channel, as expected.

Burst-error patterns, including that of Table I, were obtained from a fading simulator.[5] The main components of the simulation were a pseudorandom binary input, an FM transmitter-receiver, a Rayleigh fader that took into account desired ratios of vehicle speed to carrier wavelength, a noise generator, a pseudorandomly modulated carrier to approximate the effect of co-channel interference, and the option of switched-diversity reception. The numbers for the burst errors in Table I represent the impairment for a 24-kb/s signal-bit sequence (the bit duration determines the number of bits affected by a fade) when the mobile radio link is characterized by two-branch diversity reception under the following (worse than average) conditions:

$$\text{Signal-to-interference ratio} = 9 \text{ dB}$$
$$\text{Signal-to-noise ratio} = \infty$$
$$\frac{\text{vehicle speed}}{\text{carrier wavelength}} = \frac{V}{\lambda} = \frac{29 \text{ mi/h}}{0.353 \text{ m}} = 36.2 \text{ Hz.} \tag{1}$$

Table I — Average and median values of D and I $[P(E) = 0.03]$

Errors	I_{average}	D_{average}	I_{median}	D_{median}
Bursty	50.0	1.60	2.7	0.8
Random	32.0	1.03	23.0	0.2

2.2 Scrambled errors

Scrambled errors are of interest in sections of this paper that assume the scrambling of signal bits for error protection. The idea of bit scrambling is to expose adjacent coder bits to channel conditions that tend to be statistically independent. If the scrambling is pseudo-random, the receiver can put the received bits in proper sequence by an inverse unscrambling operation. To avoid the two operations of scrambling and unscrambling, the situation was simulated in our experiment by scrambling the known bit-error pattern and leaving the signal bits in their original sequence.

Error sequences consisted of binary entries (error bits) E that were either 0 or 1, and each entry of 1 represented a bit error in the decoding of a corresponding signal bit. The scrambling was accomplished as follows. The error-bit sequence E was handled in blocks that were M bits long, and each bit got a new position, given by a pseudo-random number (of bit intervals), as was derived from the current state of a maximal-length shift register with $\log_2 M$ stages.[6] The value of M was set at 1024, and the effect of scrambling is illustrated in Fig. 4, which is a snapshot of part of the 0.025 error-rate data. The three sections in the figure represent (contiguous) error sequences that are 1024 bits long (512 per row, two rows per block). In each of the six

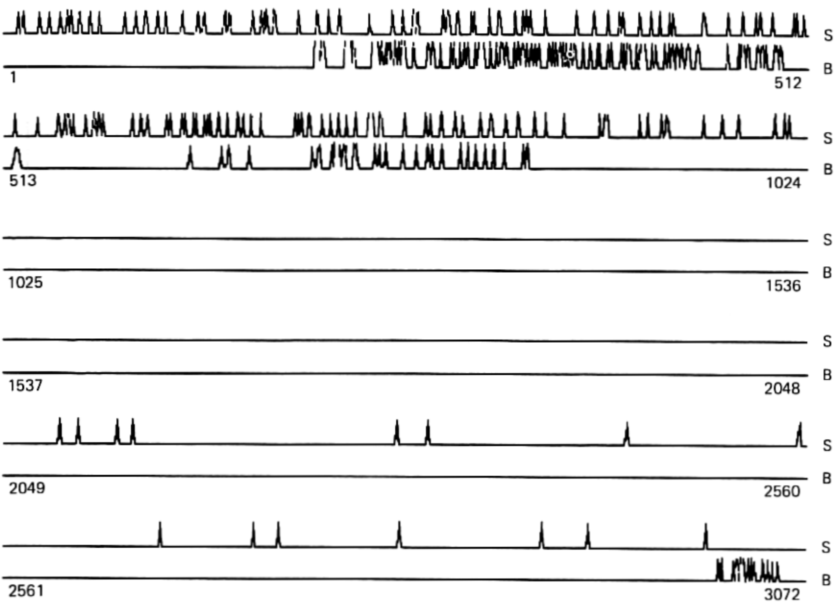

Fig. 4—Illustration of error scrambling [$P(E) = 0.025$, $V/\lambda = 36.2$ Hz].

rows, the lower sequence contains the burst errors (letter B), and the upper sequence has the scrambled errors (letter S). It is clear that a block length of $M = 1024$ is insufficient for true error randomization. Speech recordings have indicated, on the other hand, that values of M as small as 64 are sufficient to achieve useful speech encryption with signal-bit scrambling, assuming the 24-kb/s bit rate mentioned earlier.

2.3 Transmission rate and average error probability P(E)

Our simulations involved signal transmission rates of 24, 32, 40, and 48 kb/s. It is reasonable, under assumptions of constant baud rate (number of channel symbols/second), to expect higher bit-rate transmissions to be subject to correspondingly higher error rates. For example, if 24 and 48 kb/s represent two-phase and four-phase modulations of channel symbols, respectively, at a fixed 24-kHz symbol rate, the average error probability in the 48-kb/s system is expected to be typically two times* as large as that in the 24-kb/s scheme.[5] In the light of this, when we compare similar systems operating at significantly different bit rates in this paper (for example, 24 versus 48 kb/s) we assume average bit-error probabilities that are appropriately different (for example, 0.025 for 24-kb/s transmissions and 0.055 for 48-kb/s transmissions). Burst errors and scrambled errors are indicated by the notations EB and ES.

III. DM-AQF

Figure 5 illustrates the principles of variable step size delta modulation with a forward control of step size. The buffer shown in the encoder stores N input samples (typically, in linear PCM format) that are used to calculate the best step size Δ for the (future) delta modulation of the stored input block. *The step size Δ is recomputed exactly once, and explicitly transmitted to the receiver, for every block of N samples.* The rest of Fig. 5 merely represents a conventional linear delta modulator-demodulator pair.[1] The predictor is assumed to be time-invariant, and of first order. The equations describing the delta modulations are formally summarized below.

$$\begin{aligned} b_r &= \text{sgn}\,(X_r - h_1 \cdot Z_{r-1}). \\ Z_r &= h_1 \cdot Z_{r-1} + \Delta \cdot b_r. \\ Z'_r &= h_1 \cdot Z'_{r-1} + \Delta \cdot b'_r. \end{aligned} \tag{2}$$

The time indices r and $r - 1$ are not shown in the figure; however, the

* Strictly speaking, this number is a function of the carrier-to-noise ratio and the modem that is employed. For example, the number can exceed two (for a typical carrier-to-noise ratio) if FSK is used as the modulation system instead of PSK (for transmitting the speech bits over the analog channel).

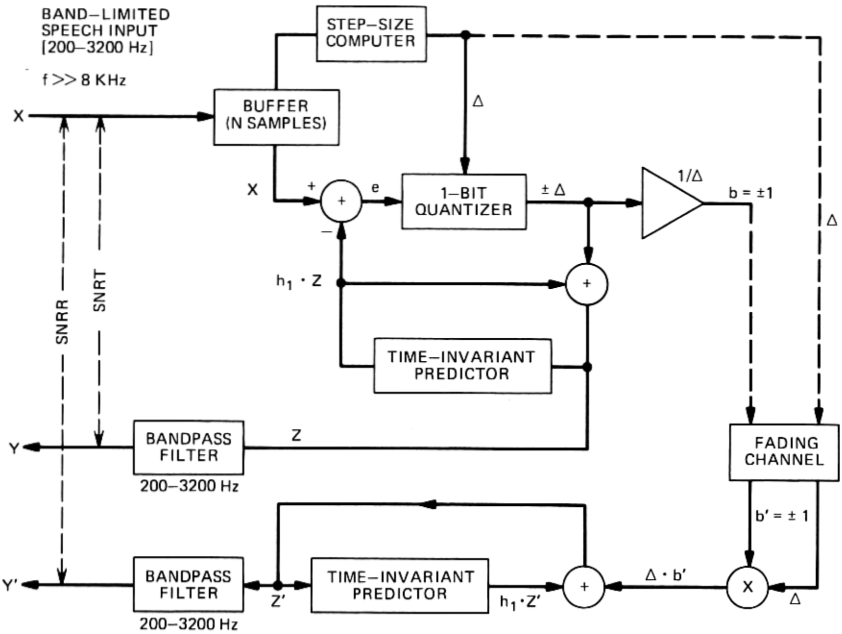

Fig. 5—Block diagram of DM-AQF codec.

implied one-sample delay occurs in the first-order predictor. Z and Z' are unfiltered staircase functions at the transmitter and receiver. The received bit b' differs from b if the error bit E is 1, and the step-size information Δ is assumed to be error-protected. For useful delta modulation, the sampling rate f should be much greater than the Nyquist frequency of the band-limited speech.

3.1 Design of Δ, N, and h_1

So that the best step-size Δ may follow the statistics of the input speech, the following algorithms were examined.

$$\Delta = K_1 \cdot \sum_{r=2}^{N} |X_r - X_{r-1}| \cdot \frac{1}{N-1}. \tag{3a}$$

$$\Delta = K_2 \cdot [\underset{2<r<N}{\mathrm{Max}} |X_r - X_{r-1}|]. \tag{3b}$$

$$\Delta = K_3 \cdot \left[\sum_{r=2}^{N} (X_r - X_{r-1})^2 \cdot \frac{1}{N-1} \right]^{\frac{1}{2}}. \tag{3c}$$

Figure 6 plots the signal-to-noise-ratio SNRT at the encoder as a function of K_n ($n = 1, 2, 3$) for the above algorithms. The numbers refer to 24-kHz delta modulation. Each scheme exhibits an optimal K_n that

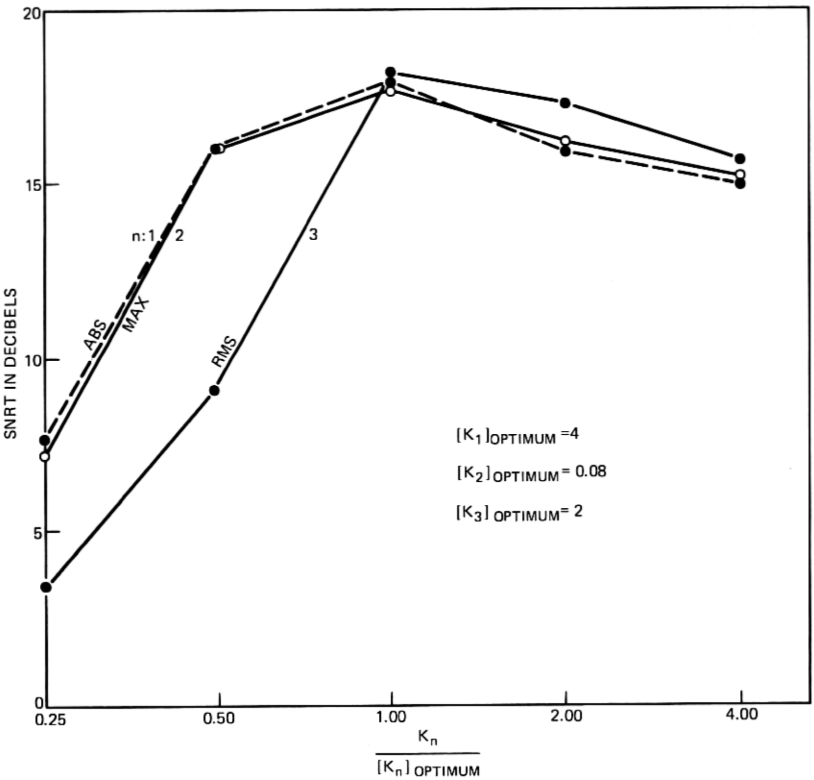

Fig. 6—Step-size computation for DM-AQF.

obviously represents the best mixture of slope-overload distortion (which predominates for $K \ll K_{OPT}$) and granular noise (which takes over for $K \gg K_{OPT}$). It is interesting that the maximum performances of the three algorithms are practically the same. This suggests that the step sizes resulting from these algorithms may not be significantly different, when optimal K values are employed. The rest of the paper will assume the use of the "average absolute slope" formula

$$\Delta = \frac{4}{N-1} \cdot \left[\sum_{r=2}^{N} |X_r - X_{r-1}| \right]. \tag{4}$$

Strictly speaking, this formula is optimal only for 24-kHz sampling and for perfect integrators ($h_1 = 1$). However, corrections for these factors were not found to be very significant for the values of f and h_1 used in our study, and formula (4) was therefore uniformly assumed for simplicity. (We may mention, however, that step-size dependencies on h_1 and bit rate will be of interest in the design of DPCM-AQF encoders; this is seen in Section IV.)

The buffer length was set at $N = 256$. This represents a compromise among three factors: (*i*) a need to minimize the encoding delay (this suggests smaller N values), (*ii*) a need to keep down the information rate in the step-size transmitting channel (this suggests slower updating, or larger N values), and (*iii*) the need to track the changing statistics of speech with an appropriate speed. A buffer length of 5 to 10 ms turns out to be a good choice for differential coding (this is demonstrated quantitatively in the context of DPCM-AQF); and $N = 256$ does indeed correspond approximately to a 10-ms delay for $f = 24$ kHz (and a 5-ms delay for $f = 48$ kHz).

The predictor coefficient was set to be $h_1 = 0.9$. This was nearly optimal from an SNRT viewpoint for the sampling frequencies of interest. Over a noisy channel, if one uses SNRR as a performance criterion, optimal values of h_1 tend to be smaller than 0.9. This is because "leakier" integrators mitigate error propagation in the output of a differential decoder. Once again, in the interest of simplicity, a quantitative consideration of this phenomenon has been deferred to the case of multibit DPCM coding (Section IV).

3.2 Bit scrambling

Table II demonstrates how bit scrambling can provide an SNRR advantage in the presence of errors. As mentioned earlier, bit scrambling was simulated by using scrambled errors ES (in place of burst errors EB for an unscrambled bit stream). Informal listening tests indicate that the perceptual advantages of bit scrambling in DM-AQF are more significant than what the SNRR gains in Table II may suggest.

3.3 Error protection by redundant coding: EP-DM-AQF

We studied a redundant DM-AQF coder in which every pair of adjacent DM bits was protected by the transmission of a (contiguous) parity check bit. When the parity failed at the receiver, a possible bit error was detected, and the received DM bit pair were forced to form an alternating (+ − or − +) sequence. This is equivalent to the

Table II — Effect of bit scrambling in DM-AQF [$P(EB) = P(ES) = 0.055$ and entries are SNRR values in dB]

f(kHz)	Speech	Burst Errors	Scrambled Errors
32	Male	7.6	8.0
40	Female	7.8	8.8

Table III — Comparison of DM-AQF and EP-DM-AQF

Scheme	f(kHz)	Transmission rate (kb/s)	$P(ES)$	SNRR (dB)
EP-DM-AQF	32	48	0.055	8.0
DM-AQF	32	32	0.055	7.1
DM-AQF	32	32	0.025	10.0

imposition of a zero-slope segment in the speech waveform when the receiver has no confidence in the incoming bits. Table III compares the performance of this error-protected system (EP-DM-AQF) with that of an unprotected DM-AQF coder, for the example of scrambled errors. The unprotected system has a bit rate of 32 kb/s, while the EP-DM-AQF operates at $32 \times \frac{3}{2} = 48$ kb/s. We are not concerned at this point with questions like a specific baud rate. However, in view of transmission rate versus error probability relations over real channels (Section II), the interesting comparison in Table III is between rows 1 and 3 (rather than between 1 and 2). It appears that the simple parity-check-based error protection is not being useful; the advantages due to error detection at the receiver are being offset (or more than offset) by the increased error probability characteristic of the higher transmission rate in EP-DM-AQF. A similar result has been obtained in a simulation of DM-AQF with correlated errors, and also with DPCM encoders where only the most significant bit is error-protected by the use of redundancy.[3]

3.4 Unprotected DM-AQF with bit scrambling; choice of f

We have considered in some detail the specific case of unprotected (nonredundant) DM-AQF with bit scrambling. Table IV presents SNRT and SNRR values for such a system at different values of f and matched values of error probability, $P(ES)$. Some entries in Table IV are interpolated values because error sequences with the corresponding $P(ES)$ values were not available. As suggested earlier in the example of binary versus quaternary PSK, an obviously meaningful comparison is between rows 1 and 4 whose error ratios differ by a factor of two.

Table IV — DM-AQF; Effect of f

f(kHz)	$P(ES)$	SNRT (dB)	SNRR (dB)
24	0.023	17.1	8.6
32	0.032	21.2	9.6
40	0.040	23.7	10.5
48	0.048	26.0	11.2

At the transmitter end, the quantization noise is easily perceived at $f = 24$ kHz. It is barely apparent at $f = 32$ kHz, and a choice of $f = 40$ kHz is likely to be more than adequate for many situations. Notice that, as f increases, so does the difference between SNRT and SNRR; and the quantization noise has a lesser and lesser influence on the speech quality at the receiver because of the relatively greater contributions of channel noise.

3.5 A comparison with syllabic-companded DM-AQB

To demonstrate that forward step-size coding is indeed desirable for the mobile radio channel, the DM-AQF scheme was compared with a syllabic-companded delta modulator with backward-step-size control (AQB). The step-size algorithm for the DM-AQB was

$$\begin{aligned} \Delta_r &= 0.966 \cdot \Delta_{r-1} + 25 \cdot [\text{ADAPT}]_r \\ [\text{ADAPT}]_r &= 1 \text{ if } \left| \sum_{s=0}^{3} b_{r-s-p} \right| = 4, \text{ for } p = 0 \text{ or } 1 \text{ or } 2 \\ &= 0 \text{ otherwise.} \end{aligned} \quad (5)$$

The algorithm is reminiscent of, if not identical to, the digitally controlled delta modulation (DCDM) scheme due to Greefkes,[7] which is an AQB technique well known for its error resistance. Figure 7 demonstrates that, in the presence of bit errors, the performance of DM-AQF degrades more gracefully than that of the DM-AQB defined in (5). It must be remembered, of course, that the DM-AQB system is implemented more easily and without encoding delay.[8]

3.6 The problem of step-size transmission in AQF

We have tacitly assumed so far that step-size information in DM-AQF can be very reliably transmitted, even over a fading channel, because step-size updating has to be done only infrequently. We shall now demonstrate this with some numbers.

Figure 8 illustrates a histogram of step sizes that resulted from utilizing (4) for a 32-kHz DM-AQF encoder. It was noted that the encoding was very tolerant to a maximum step-size constraint of 155, and a step-size resolution equal to 10; in other words, to a step-size dictionary of only 16 steps (5, 15, $\cdots$, 155). In practice, the maximum-to-minimum step size ratio would probably be greater than 31, in anticipation of highly nonstationary speech inputs.

The four-bit step-size information was transmitted as follows. At the beginning of each block of $N = 256$ bits, the respective four-bit word was transmitted five consecutive times. Each bit in the step-size word was decoded on the basis of a majority count over the five

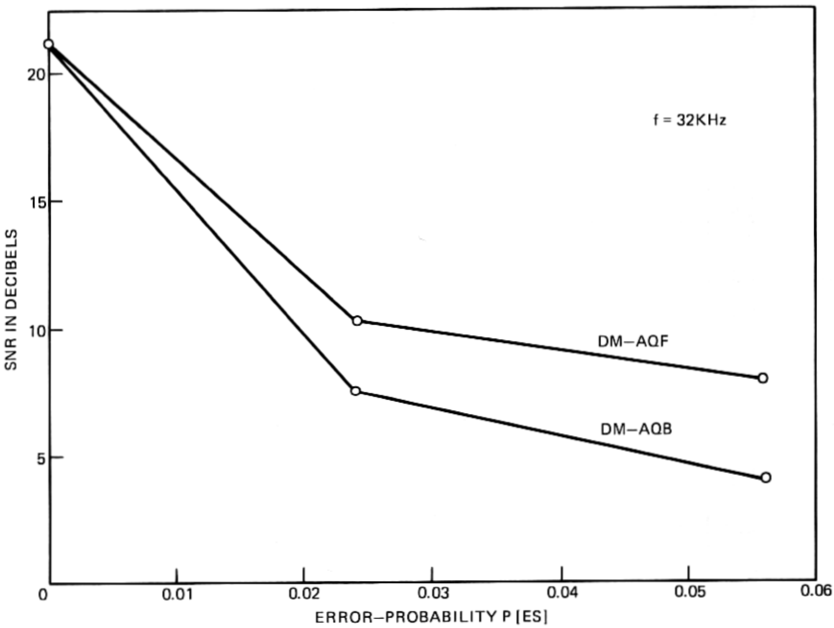

Fig. 7—Comparison of DM-AQF and DM-AQB.

received versions of the bit. The step-size transmissions increased the overall bit rate from 32 kb/s to $32[(256 + 4 \times 5)/256] = 34.5$ kb/s. For a random error rate of $P(ES) = 0.025$, the SNRR with the explicit transmission of step size, as above, was nearly identical with the value obtained in a simulation that tacitly assumed the presence of correct step-size information at the receiver. The result is not surprising; the probability of failure of a majority count of order 5 is given by

$$P(\text{M.C}; 5) = \sum_{r=3}^{5} p^r (1-p)^{5-r} \binom{5}{r}$$
$$\sim 10p^3 \quad \text{if} \quad p \ll 1, \tag{6}$$

where p is the error probability. With $p = P(ES) = 0.025$, $P(\text{M.C}; 5) = 1.64 \times 10^{-4}$. The probability that at least one of the bits of a step-size word is wrongly decoded in our scheme is therefore no greater than 6.4 in 10,000, and there were only 250 step-size transmissions during the entire length of the (2-s) speech utterance being coded.

3.7 SNRT, SNRR, and P(E) as functions of time

We conclude our discussion of DM-AQF with an interesting demonstration of the time dependencies of SNRT, SNRR, and $P(ES)$, as measured over blocks that were $N = 256$ samples long. The sampling rate was

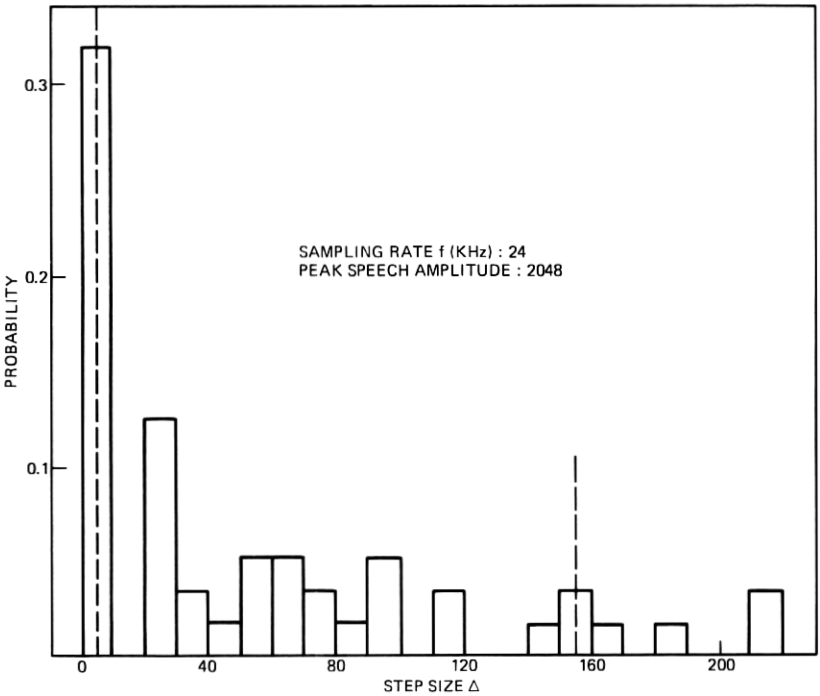

Fig. 8—Histogram of step sizes in DM-AQF (peak speech amplitude = 2048).

24 kHz, the average error probability was 0.025 [refer to Fig. 2 and eq. (1) for error characterization], and the plots on Fig. 9 used numbers taken once every 20 blocks (5120 samples). Notice the obvious negative correlation between the time functions SNRR$[t]$ and $P(ES)[t]$. The time variation of SNRT is, of course, purely a reflection of the input speech material.

IV. DPCM-AQF

Figure 10 is a block diagram of differential PCM with forward step-size control. Differences from Fig. 5 consist in the use of a B-bit quantizer ($B = 3$ or 4 in this paper), and in the assumption of Nyquist-rate sampling, which obviates the need for a critical output filter. Basic DPCM notation is as follows: W is the normalized code word magnitude, e is the prediction error, and $\tilde{e}$ is the quantized value of e. The time-invariant (first-order) predictor coefficient is h_1, and r represents an instantaneous (sampled) value. The received bits b_q' ($q = 1, 2, \cdots, B$) are different from the transmitted bits b_q if a corresponding error bit E equals 1. The step size is Δ; it is assumed to be recalculated once every N samples, and successfully error-protected in transmission. The

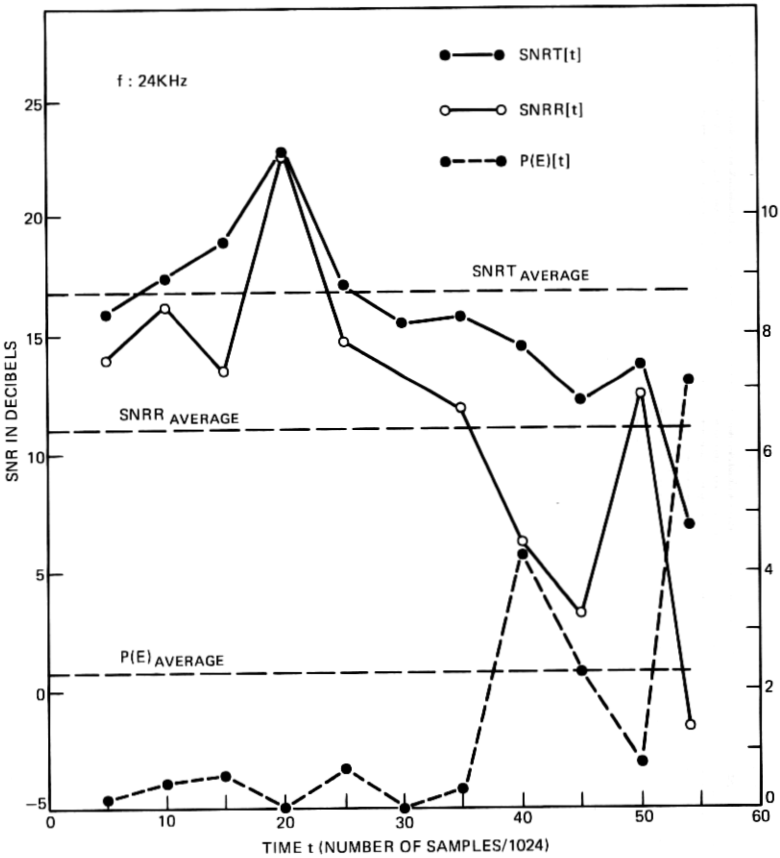

Fig. 9—Time variation of SNRT, SNRR, and $P(E)$ (V/λ = 36.2 Hz).

following are the salient DPCM equations.

$$
\begin{aligned}
b_{qr} &= \pm 1; \quad q = 1, 2, \cdots, B. \\
bb_{qr} &= 0.5 b_{qr} + 0.5 = 0 \text{ or } 1. \\
e_r &= X_r - h_1 \cdot Y_{r-1}. \\
Y_r &= h_1 \cdot Y_{r-1} + \tilde{e}_r. \\
Y'_r &= h_1 \cdot Y'_{r-1} + \tilde{e}'_r. \\
\tilde{e}_r &= W_r \cdot \Delta. \\
W_r &= \left[\sum_{q=2}^{B} 2^{B-q} \cdot bb_{qr} \right] \cdot \operatorname{sgn} b_{1r}.
\end{aligned} \tag{7}
$$

For any sample r, the sign of the code word W is the most significant bit b_1; the least significant bit is b_B.

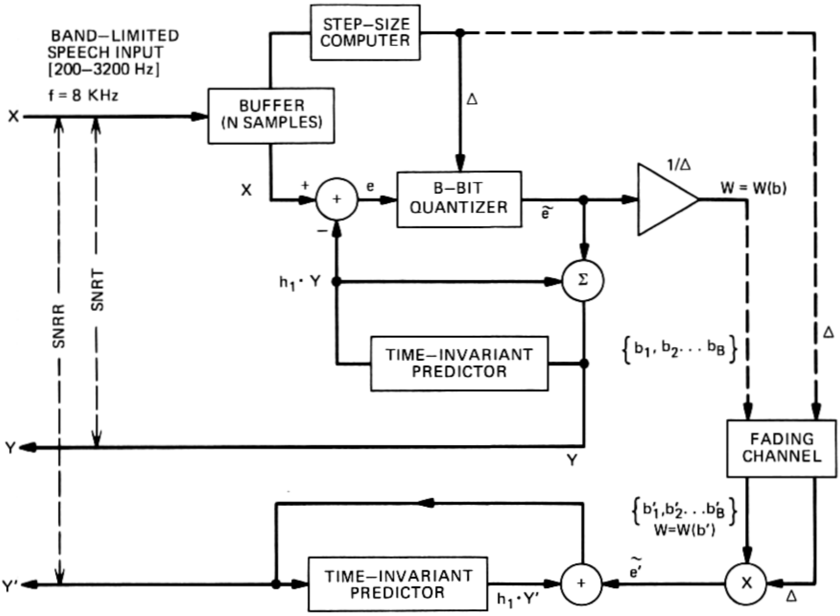

Fig. 10—Block diagram of DPCM-AQF codec.

4.1 Design of Δ, N, and h_1

AQF step sizes are derived (once for every block of N samples), using the formula

$$\Delta = K_4 \cdot \frac{1}{N-1} \cdot \sum_{r=2}^{N} |X_r - h_1 \cdot X_{r-1}|. \tag{8}$$

Figures 11, 12, and 13 illustrate typical SNRT and SNRR dependencies on the parameters K_4, N, and h_1, respectively. The curves refer to the case of $B = 4$, $P(EB) = 0.055$, and to a redundant transmission technique described in Fig. 14. It is clear that SNRT-maximizing designs are significantly different from the SNRR-maximizing values. Rather than getting bogged down in the controversial question of whether SNRT or SNRR is to be used as a performance criterion, we have elected, arbitrarily, to discuss the following SNRR-maximizing designs that were approximately good for the $P(E)$ range of 0.025 to 0.055:

$$\begin{aligned} N &= 64 \\ h_1 &= 0.6 \\ K_4 &= 0.50 \quad \text{if} \quad B = 3 \\ &= 0.25 \quad \text{if} \quad B = 4. \end{aligned} \tag{9}$$

Notice that, in Fig. 12, SNRR is maximum at $N = 128$. However, the

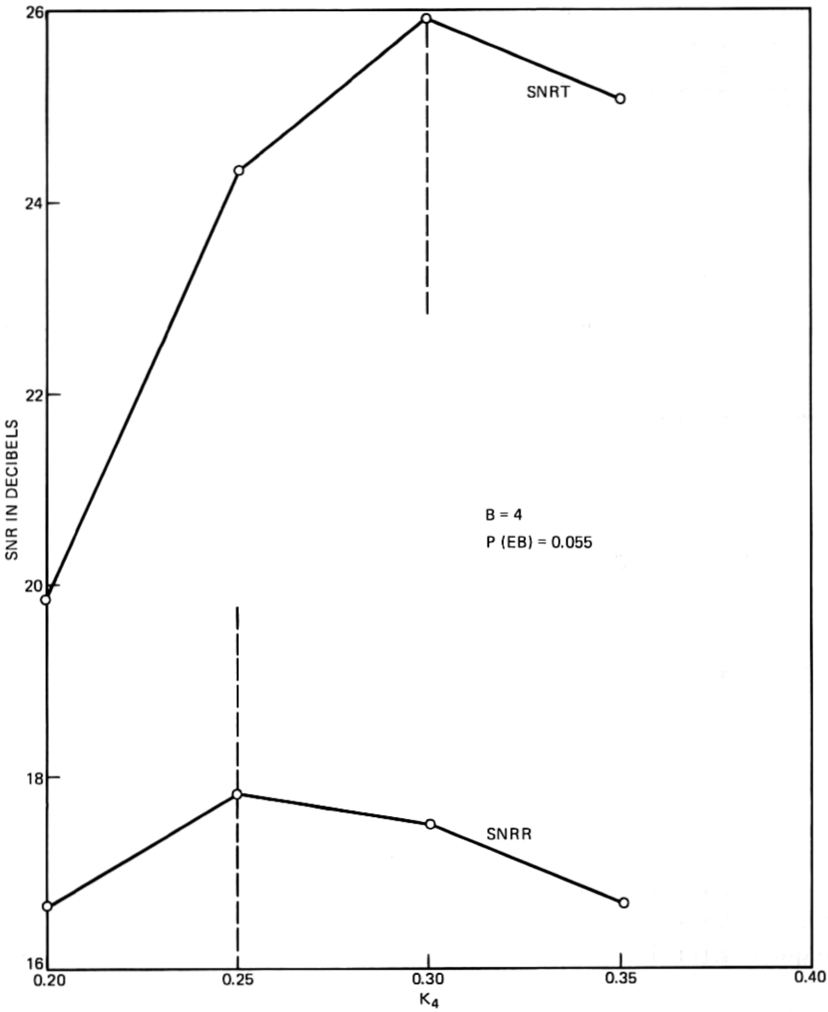

Fig. 11—Step-size computation for DPCM-AQF.

encoding delay is less objectionable (8 ms, instead of 16 ms) with $N = 64$. Note also that SNRT-maximizing designs call for higher values of both h_1 and K_4.

The maximum-to-minimum step-size ratio in the simulation was about 1000. It is possible to reduce this ratio to 100, and still provide useful coding of nonstationary speech.[1] Smaller step-size ratios enhance bit-error resistance. They also tend to simplify the problem of transmitting step-size information.

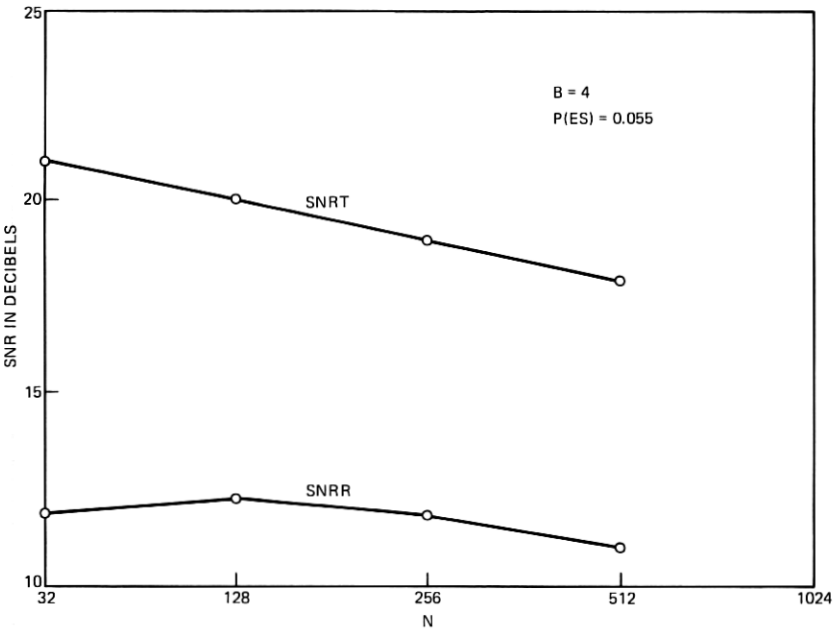

Fig. 12—Step-size updating in DPCM-AQF.

4.2 Error-protected DPCM-AQF

Figure 14 illustrates the use of time-diversity coding designed to protect DPCM bits from burst errors. The time-diversity is provided by the delay P that will be discussed presently. Figure 14a defines a three-bit EP-DPCM system where the most significant bit b_1 is transmitted three times, and the second most significant bit b_2 is sent twice. The least significant bit b_3 is transmitted only once. At the receiving end, the value of b_1 is determined on the basis of a majority count over the three received versions. In regard to the magnitude bit b_2, if the two versions of b_2 do not agree, the receiver code word is forced to its smallest magnitude (the polarity is still defined by the unequivocally decoded value of b_1). This is equivalent to forcing a "minimal-slope" segment in the decoded speech waveform when the receiver is in doubt about the code-word magnitude. Figure 14b defines a four-bit EP-DPCM system where only the most significant bit b_1 is error-protected. Once again, the decoding of b_1 at the receiver follows a majority count over the three received versions thereof. Assuming 8-kHz sampling, both EP-DPCM systems of Fig. 14 would operate at 48 kb/s. However, the three-bit system of Fig. 14a has the benefit of greater error protection.

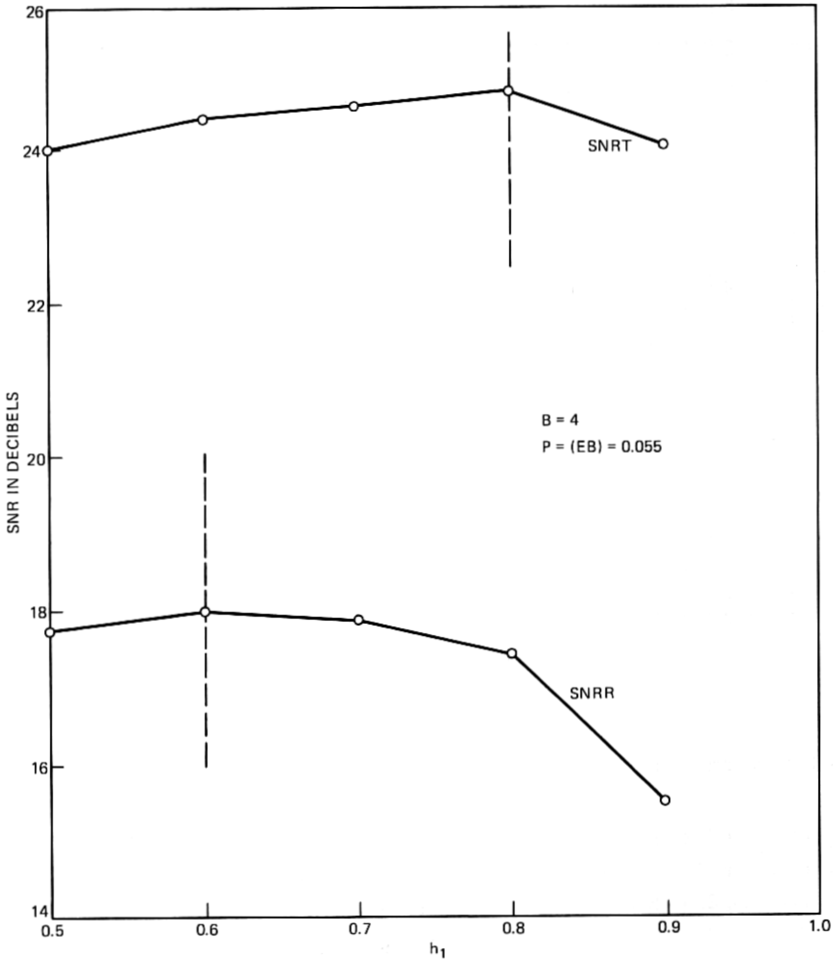

Fig. 13—Design of predictor coefficient h_1 in DPCM-AQF.

Figure 15 shows the benefits of time diversity for the example of the four-bit system of Fig. 14b. It is interesting that SNRR is still tending to increase at P values as large as 1024. It can be expected that, if $P \gg [D + I]_{\text{average}}$, successive repetitions of a given bit tend to be affected independently by the channel. D and I are the burst duration and spacing mentioned in Section II. The DPCM-AQF coders of this paper assume a uniform value of $P = 768$. For a bit rate of 48 kb/s, this implies a total encoding delay (from Fig. 13) of 2 P bits, or about 32 ms.

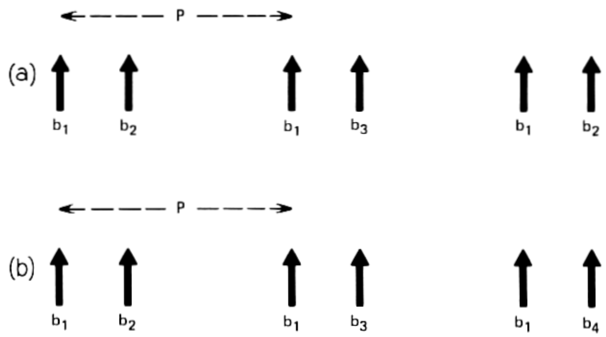

Fig. 14—Time-diversity coding for EP-DPCM.

4.3 EP-DPCM; choice of B

We now compare the two 48-kb/s systems of Fig. 14. Table V shows, for two different speech inputs, the SNRR values obtained with the three-bit and four-bit systems. The greater error protection in the three-bit system seems to make it more robust, in spite of the better quantization noise (SNRT) performance of a four-bit coder, and the better receiver-end quality of three-bit coded speech is very obvious in listening tests. The result is also mentioned by Noll.[3] It is true that four-bit coding can provide a 6-dB superiority in SNRT. It appears, on the other hand, that the subjective SNRT in DPCM is known to be considerably higher than a measured objective SNRT,[1] and the SNRT of

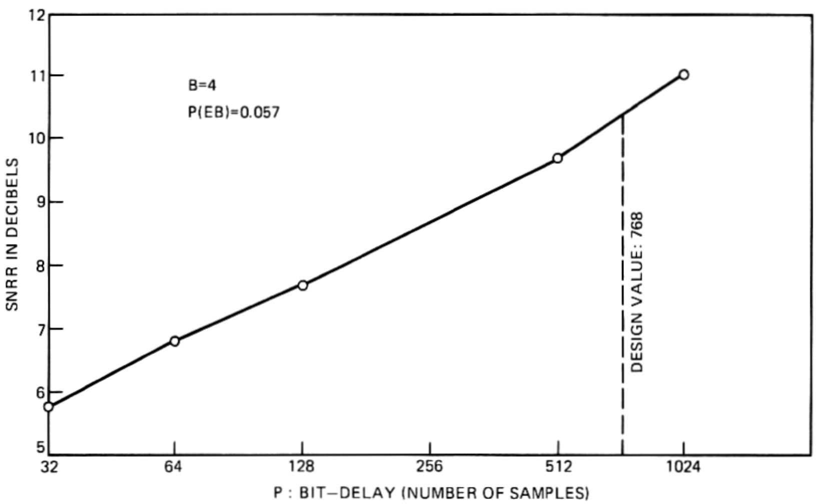

Fig. 15—Effect of time diversity on received-speech quality ($V/\lambda = 36.2$ Hz).

Table V — EP-DPCM; Comparison of three- and four-bit systems (entries are SNRR values in dB; rows 1 and 2 represent different speech inputs)

$P(EB)$	$B = 3$	$B = 4$
0.025	14.2	12.8
0.055	9.7	9.3

three-bit DPCM may prove to be subjectively adequate for some mobile links. In that case, the system of Fig. 14a would be a good configuration for error-protected DPCM.

Table VI further demonstrates the benefits of error protection for $B = 3$. In view of the transmission rate—$P(EB)$ relationships mentioned in Section II, the interesting comparison in the table is between rows 1 and 3, not 1 and 2. It is seen that EP-DPCM at 48-kb/s provides a better SNRR than unprotected 24-kb/s DPCM, in spite of the higher error probabilities that accompany the 48-kb/s transmissions. This contrasts interestingly with the results of Table II where error protection was seen to be ineffective for DM-AQF. The suitability of error protection for DPCM (and not DM) seems to be a direct consequence of the multibit quantization in DPCM: it is possible to isolate and error-protect only the more significant DPCM bits and incur an overall redundancy of 50 to 100 percent; a majority count for a 24-kHz DM would immediately result in a transmission rate of 72 kb/s (and a redundancy of 200 percent).

In a recently proposed, and not less effective, approach to EP-DPCM coding,[9] the DPCM bits are error-protected in suitably long blocks rather than on a bit-by-bit basis: The time diversity reception consists in selecting one of two time-separated blocks on the basis of an autocorrelation-type quality evaluation at the receiver.

4.4 Bit scrambling in DPCM

Informal listening tests, as well as SNRR evaluations, have shown that bit scrambling, and the resulting error-randomization, is much less effective for multibit DPCM than for DM. The reason for this is not

Table VI — Benefits of error protection for DPCM ($B = 3$)

Code	Transmission Rate (kb/s)	$P(EB)$	SNRT	SNRR
EP-DPCM	48	0.055	19.4	12.4
Unprotected DPCM	24	0.055	19.4	7.1
Unprotected DPCM	24	0.025	19.4	9.6

well understood. However, situations exist where bit scrambling can provide nominal SNRR gains even for DPCM. These have been noted by Noll.[3]

4.5 DPCM-AQB

The problem of step-size transmission for DPCM-AQF is expected to be handled through techniques not very different from those discussed in the context of DM-AQF. Following the calculation procedures of that section, it is estimated that virtually error-free transmission of DPCM step size would be possible (for the error rates considered) by the expenditure of about 5 kb/s of channel capacity.* To indicate the desirability of dedicating this kind of channel capacity for step size, we investigated two types of backward step-size control. One of these was adaptive differential quantization with a one-word memory.[1] Here, the quantized step size is modified for every sample by a factor determined solely by the magnitude of the latest code word W_r. The other adaptive scheme derived step-size information by an algorithm similar to the DPCM-AQF rule (8). The summation, however, was over the most recent N samples of quantized speech. Neither of the above backward schemes performed well enough with bit errors to merit inclusion of their results. It is conceivable, however, that, as in DM, some kind of a slowly adapting or syllabic DPCM may provide a fair result for mobile radio. It is also conjectured that the performance of such a scheme would be upper-bounded by that of DPCM-AQF in the manner of Fig. 7. At least one approach to slowly companded DPCM has been proposed to date.[10,11]

V. CONCLUSION

The object of this paper was to specify two differential coders—one from the DM family and the other from the DPCM class—that would be appropriate for digitizing speech in some types of mobile radio systems. The results of our work indeed suggest two such coders: a nonredundant 40-kHz DM-AQF coder with bit scrambling and an error-protected three-bit DPCM-AQF operating at a nominal 48 kb/s. The typical capabilities of these systems are summarized in Table VII, which is based on the example of a female utterance, "The lathe is a big tool." The transmission rates and error probabilities in Table VII are matched, albeit in a limited sense, as discussed earlier. Also, as emphasized already, the error rates in Table VII are worse-than-average numbers for many mobile radio links.

* If the overall transmission rate of the system is constrained to be 48 kb/s, it may be possible to work with a sampling rate of about 7 kHz, instead of 8 kHz, to accommodate the step-size information in the 48-kb/s channel (Ref. 4).

Table VII — Comparison of DM-AQF and EP-DPCM-AQF

Coder	B	f(kHz)	Transmission rate kb/s	Estimate of δ(kb/s)	Bit-Error Probability	SNRT (dB)	SNRR (dB)
EP-DPCM-AQF	3	8	$48 + \delta$	5	$P(EB) = 0.055$	20.5	14.5
DM-AQF	1	40	$40 + \delta$	<5	$P(ES) = 0.045$	23.7	10.1

In assessing the coders of Table VII, it may be worth noting that the DM system is more flexible. For example, the DM sampling rate can be lowered to 32 kHz with only a 2.5-dB loss in maximum speech quality SNRT (Table IV). Further, if the refinements of time-diversity coding (in DPCM) and bit scrambling (in DM) are eschewed, it is our experience that the DM system will lose less in the process.

Obviously, a common denominator in the above systems is adaptive differential quantization. Crudely speaking, adaptive quantization serves to squelch channel noise, while differential coding tends to smear it; and the combination appears to be perceptually very desirable in the context of mobile telephony.

Formal perceptual studies in this subject should appropriately include other digital techniques such as nondifferential (PCM) and backward-adaptive (AQB) coders. The studies should also include the possible effects of encoding delay. Clearly, the amount of this delay depends on what combination of refinements (forward coding, bit scrambling, and time diversity) is employed; and if the total delay gets to be long enough, the benefits of a better SNRR (due to reliable step-size information, error randomization, and redundant error protection, respectively) may be accompanied by a loss of echo performance over certain kinds of networks. The best compromise between transmitted speech quality, received speech quality, and encoding delay is very likely to be system-specific; and the nature of this compromise may influence or define a selection among analog techniques, conventional digital schemes (AQB), and step-size transmitting codes (AQF).

VI. ACKNOWLEDGMENTS

The idea of forward quantization for the fading channel was first suggested by P. Noll who, together with R. W. Schafer and M. R. Karim, provided an extensive amount of stimulation to this work. Error tapes were provided by V. H. MacDonald and G. Arredondo. Several very helpful comments on an earlier draft of this paper came from G. Arredondo and B. H. Bharucha.

REFERENCES

1. N. S. Jayant, "Digital Coding of Speech Waveforms—PCM, DPCM, and DM Quantizers," Proc. IEEE, May 1974, pp. 611–632.
2. G. A. Arredondo and J. I. Smith, "Voice and Digital Transmission in a Mobile Radio Channel at 850 MHz," Proc. Nat. Elec. Conf., *29*, 1974, pp. 74–79.
3. P. Noll, "Effects of Channel Errors on the Signal-to-Noise Performance of Speech Encoding Schemes," B.S.T.J., this issue, pp. 1615–1636.
4. D. L. Dutweiler and D. G. Messerschmidt, "Experimental Mobile Radio Digital Transmission with Time Diversity," Proc. Inter. Commun. Conf., San Francisco, June 1975.
5. V. H. MacDonald and G. Arredondo, private communication.
6. R. G. Gallager, *Information Theory and Reliable Communication*, Sec. 6.6, New York: John Wiley, 1968.
7. J. A. Greefkes, "Code Modulation with Digitally Controlled Companding for Speech Transmission," Philips Tech. Review, *31*, No. 11/12, pp. 335–353, 1970.
8. M. R. Karim, "Delta Modulation of Speech for Mobile Radio," unpublished work.
9. N. S. Jayant, "An Autocorrelation Criterion for the Time Diversity Reception of Speech Over Burst Error Channels," B.S.T.J., this issue, pp. 1583–1595.
10. S. U. H. Qureshi and G. David Forney, Jr., "A 9.6/16 KBPS Speech Digitizer," Proc. Int. Commun. Conf., San Francisco, June 1975.
11. D. J. Goodman, "A Robust Adaptive Quantizer," to be published in IEEE Trans. on Commun.