AT&T BELL LABORATORIES TECHNICAL JOURNAL
Vol. 63, No. 4, April 1983
Printed in U.S.A.

Weighting Strategies for Companded PCM Transmitted Over Rayleigh Fading and Gaussian Channels

By C.-E. SUNDBERG,* W. C. WONG,† and R. STEELE‡

(Manuscript received July 29, 1983)

We present three methods of weighting N-bit μ-law Pulse Code Modulation (PCM) with binary modulation, and evaluate their performances theoretically for transmission over Rayleigh fading and Gaussian channels. The digital modulation methods considered are noncoherent frequency shift keying and coherent phase shift keying. Ideal selection combining and ideal maximal ratio combining diversity techniques are employed when the transmission is over fading channels. Weighting System 1 is the conventional weighted pulse code modulated system, where the bits in every PCM word have the same weighting profile. Weighting System 2 has 2^N unique weighting profiles, while System 3 has an addendum to System 2, where every bit in a particular word is further weighted by a unique multiplicative factor. For Rayleigh fading channels and the encoder operating at an input level that provided maximum signal-to-quantization noise-ratio, we obtained gains in overall signal-to-noise ratio (s/n) over unweighted μ-law PCM of 3, 4.5, and 6 dB for Systems 1, 2, and 3, respectively. When the systems were used in conjunction with Gaussian channels, the corresponding gains were 10, 12, and 17 dB, for a channel s/n of 10 dB. In addition to the theoretical results, we conducted computer simulations using four concatenated speech sentences transmitted via our weighted μ-law PCM systems over mobile radio channels. The simulation performances were in good agreement with our theoretical results, which were based on input signals having an exponential distribution.

* University of Lund, Sweden. † AT&T Bell Laboratories. ‡ University of Southampton, England.

I. INTRODUCTION

When analog signals are PCM-encoded into binary words, the bits in any word represent different contributions to the decoded analog sample.[1-3] For example, in 8-bit piecewise μ-law Pulse Code Modulation (PCM),[2] the first bit in each PCM word represents the polarity of the speech sample, the next three bits identify the segment number, and the final four bits locate the value of the sample within the segment. Clearly, there is a hierarchical structure, with the most significant magnitude bit being the most influential in determining the accuracy of the recovered speech sample, and the importance of the subsequent bits declining rapidly until the least significant bit (LSB) is reached.

Binary errors inevitably will occur to some extent in the regeneration process at the receiver, and therefore it seems reasonable to mitigate the effect of these errors by matching the energy of each bit prior to *transmission in a manner dependent on their contribution to the* accuracy of the recovered decoded sample. If this approach is adopted, the average energy used in the transmission of each binary word is arranged to be identical to that in conventional PCM, where all bits have the same energy. Thus the energies allocated to the most significant magnitude bits are enhanced at the expense of the least significant bits whose energies are curtailed.

This concept, due to Bedrosian,[1] has an instant appeal as a technique for reducing the noise due to binary transmission errors in the recovered speech signal. A PCM system that weights the binary levels by differing amounts, i.e., the binary amplitudes may be different for each bit in the PCM word, is known as weighted PCM. Sundberg[2] generalized the analysis of Bedrosian to include companded PCM, making use of the A factors, which he and Rydbeck conceived.[3,4]

For transmission of unweighted PCM by means of a binary modulation scheme, the bit error probability is the same for all bits, independent of their position in the word. This is clearly not so for weighted PCM, where the weighting of the bit energies reduces the error probabilities of the more significant bits. As the average energy (or power) per transmitted bit is constant, irrespective of whether weighting is applied, it follows that the error probabilities of the less significant bits must increase. However, the effect of weighting is to improve the overall signal-to-noise ratio (s/n) for a given channel s/n.

Our intention in this discourse is to describe new weighting strategies for μ-law PCM transmitted over Rayleigh fading and Gaussian channels by means of either binary Noncoherent Frequency Shift Keying (NCFSK) or binary Coherent Phase Shift Keying (CPSK) modulation. For each weighting profile and channel we will present formulae for the noise power in the recovered signal due to the effect of transmission errors. The theoretical overall s/n is then presented as a function of channel s/n and compared to simulated results using speech signals. However, before we introduce our new weighting methods we will briefly review the basic concepts of how transmission errors in companded PCM manifest as noise in the recovered analog signal.

II. DIGITAL NOISE IN CONVENTIONAL COMPANDED PCM

In companded PCM the input sample x may be initially compressed to

$$y = f(x), \tag{1}$$

where a common choice of $f(x)$ is the μ- or A-law.[3] These laws are defined as follows. For μ-law PCM,

$$f(x) = \begin{cases} \dfrac{\log(1 + \mu x)}{\log(1 + \mu)}; & 0 \leqslant x \leqslant 1 \\ -f(-x); & -1 \leqslant x < 0, \end{cases} \tag{2}$$

and for A-law,

$$f(x) = \begin{cases} \dfrac{Ax}{1 + \log A}; & 0 \leqslant x \leqslant \dfrac{1}{A} \\ \dfrac{1 + \log(Ax)}{1 + \log A}; & \dfrac{1}{A} < x \leqslant 1 \\ -f(-x); & -1 \leqslant x < 0. \end{cases} \tag{3}$$

In stating $f(x)$ we have followed the practice[2,5] of normalizing the range of both x and y to the interval -1 to $+1$. The compressed sample y is quantized to y_i, which can have one of 2^N possible values. Encoding of y_i into an N-bit binary word L_i ensues, and the words are generated at a rate in excess of the Nyquist rate. The stream of binary words is suitably filtered and modulates a carrier for transmission. Figure 1 is a block diagram of a companded PCM system, although the input filter and sampler, and the final interpolating filter are not displayed.

The receiver demodulates the incoming signal and regenerates the binary words. For the transmitted word L_i, the regenerated word is $L_{i,l}$, where the subscript l signifies that one or more bits may be erroneously produced. We may represent $L_{i,l}$ as the exclusive-OR operation on the binary vectors L_i and e_l, where e_l is an N-bit sequence whose subscript l identifies the error sequence. Thus, we may express $L_{i,l}$ as

$$L_{i,l} = L_i \oplus e_l. \tag{4}$$

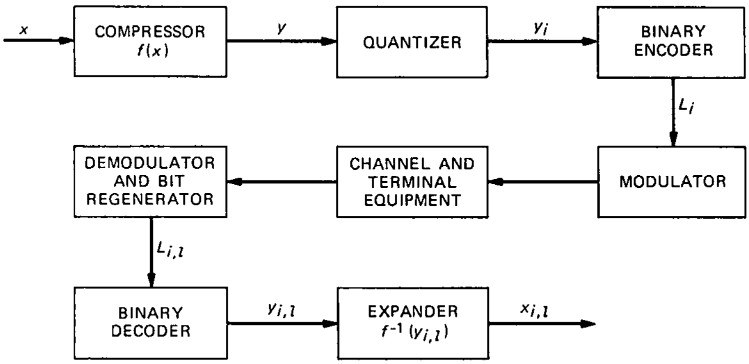


Fig. 1—Block diagram of a conventional PCM system (filters and sampler not shown).

For example, the error sequence 00110000 means that the bits in the third and fourth bit positions are erroneously regenerated. Observe that there are $2^N - 1$ possible e_l sequences. We do not consider the all-zero e_l sequence, as it implies no bit errors are present in the regenerated word. The word $L_{i,l}$ is binary decoded to $y_{i,l}$ and the recovered sample is

$$x_{i,l} = f^{-1}(y_{i,l}). \tag{5}$$

If no transmission binary errors occur, the recovered sample is

$$x_{i,l} = f^{-1}(y_i) = x_i \tag{6}$$

and the overall noise power reduces to the encoder noise power

$$\epsilon_e^2 = E\{(x - x_i)^2\}, \tag{7}$$

where the expected value is formed over the source statistics. However, we are not concerned here with the noise power generated in the encoder, but with the noise power in the recovered samples due to transmission errors, which we call digital noise power. This power is

$$\epsilon_a^2 = E_{i,l}\{(x_i - x_{i,l})^2\}, \tag{8}$$

where the expectation is performed over all 2^N levels of x_i, and over all $2^N - 1$ possible error sequences e_l. The set of $2^N - 1$ error sequences e_l may be subdivided into groups containing single, double, $\cdots$, N-bit errors per PCM word. These groups are described by the Hamming weight, w, so that $w = 1, 2, \cdots, N$. The input signal and PCM encoder are independent of the channel imperfections enabling us to express ϵ_a^2 as

$$\epsilon_a^2 = \sum_{l=1}^{2^N-1} \rho_l E_i\{(x_i - x_{i,l})^2\}$$
$$= \sum_{l=1}^{2^N-1} \rho_l A_l, \tag{9}$$

where

$$A_l \triangleq E_i\{(x_i - x_{i,l})^2\} \tag{10}$$

are the A-factors, and ρ_l is the probability of occurrence of the lth error sequence e_l, corresponding to the A-factor A_l. Observe that A_l is the average noise power in the output sequence due to the existence of the specific error sequence e_l, the expectation being made over all the values i, namely over the 2^N quantized levels.

When independent errors occur with average bit error probability P, then the probability of w errors in an N-bit word is

$$P_w = P^w(1 - P)^{N-w}.$$

Thus, when the error sequence e_l has w bit errors in the N-bit word, $P_l = P_w$. A case of particular interest is when no more than one error occurs in any PCM word and $P \ll 1$, whence $P_w \simeq P$, and the digital noise power becomes

$$\epsilon_a^2 = P \sum_{l=1}^{N} A_l,$$

where $A_1, A_2, \cdots, A_N$, are the single-error A-factors for errors in bits 1 to N.

III. WEIGHTED PCM

The digital noise power ϵ_a^2 in a PCM system is determined using eq. (9), and we now seek to weight all the bits in the transmitted signal in order to reduce ϵ_a^2. We consider the case when single bit errors occur in any PCM word, knowing that for more severe error conditions the overall s/n is, in general, unacceptably low. Thus, for this case we approximate eq. (9) as

$$\epsilon_a^2 = \sum_{l=1}^{N} \rho_l \mathrm{A}_l. \tag{11}$$

We emphasize that consideration will be given only to binary modulation schemes. From eq. (10),

$$A_l = \sum_{i=0}^{2^N-1} p_X(x_i)\Delta_i(x_i - x_{i,l})^2, \tag{12}$$

where $x_{i,l}$ and x_i are given by eqs. (5) and (6), respectively,

$$\Delta_i = \frac{2^{1-N}}{f'(x_i)} \tag{13}$$

is the quantization interval associated with x_i, $p_X(x_i)$ is the PDF of the input sequence, assumed constant over the interval Δ_i, and $f'(x_i)$ is the derivative of $f(x)$ at $x = x_i$.

The validity for this assumption is that there are many quantization levels, as indeed there are for our 8-bit μ-law PCM encoding considered in this paper. In addition to our 256 quantization levels, we have normalized the ranges of x and y to -1, $+1$, and the compressed signal y is uniformly quantized.[3] Writing

$$a_l(x_i) = (x_i - x_{i,l})^2, \tag{14}$$

and observing that the probability of x being located in the quantization interval Δ_i is

$$p_i = p_X(x_i)\Delta_i \tag{15}$$

allows us[5] to express the A-factors as

$$A_l = \sum_{i=0}^{2^N-1} p_i a_l(x_i). \tag{16}$$

Thus, the digital noise power is from eqs. (11) and (16):

$$\epsilon_a^2 = \sum_{l=1}^{N} \sum_{i=0}^{2^N-1} \rho_l p_i a_l(x_i). \tag{17}$$

The probability ρ_l of the error sequence e_l occurring depends upon the modulation scheme employed, e.g., CPSK, NCFSK, whether diversity reception is used, the channel type, the channel s/n, and the method of weighting the bits in the PCM words. We will describe three weighting schemes, which for brevity we will designate Systems 1, 2, and 3.

3.1 Weighted PCM System 1

As a means of introducing weighted PCM we will consider binary amplitude modulation. For this situation each unweighted bit b_l in every PCM word has its magnitude multiplied by $\sqrt{\phi_l}$, where $l = 1, 2, \cdots, N$ refers to the first, second, $\cdots$, LSB, respectively, and $\{\phi_l\}_{l=1}^{N}$ is the weighting profile for the PCM word, generated as l is varied from 1 to N. For a particular bit, the lth, say, ϕ_l is also referred to as its weight. Observe that in this weighting scheme, each of the 2^N words has identical weighting profiles. The N weighting samples $\sqrt{\phi_l}$, $l = 1, 2, \cdots, N$, are stored in a circulating shift register. As each PCM bit

b_l, $l = 1, 2, \cdots, N$ is generated, the appropriate value of $\sqrt{\phi_l}$ is removed from the shift register and is multiplied by b_l. Thus the weighted bit applied to the transmission terminal equipment is

$$B_{1,l} = \sqrt{\phi_l} b_l; \qquad b_l = \pm 1, \tag{18}$$

where the subscript 1 in $B_{1,l}$ signifies System 1. The receiver is identical to that of unweighted PCM, and the received bits are regenerated by observing at a sampling instant if

$$\begin{aligned} \hat{B}_{1,l} \geqslant 0; \qquad & \text{bit of logical 1} \\ \hat{B}_{1,l} < 0; \qquad & \text{bit of logical 0,} \end{aligned} \tag{19}$$

where $\hat{B}_{1,l}$ is the received value of $B_{1,l}$.

For the general modulation case, let the energy of the lth bit be

$$E_l = \phi_l E, \tag{20}$$

where E is the average energy per bit of the PCM signal,

$$E = \frac{1}{N} \sum_{l=1}^{N} E_l, \tag{21}$$

and by definition ϕ_l is the energy weight assigned to the lth bit. The constraint on the weights in System 1 is, with the aid of eqs. (20) and (21),

$$N = \sum_{l=1}^{N} \phi_l. \tag{22}$$

Thus, from eq. (20), the channel s/n for the lth bit is

$$\Gamma_l = \phi_l \Gamma; \qquad \text{System 1,} \tag{23}$$

where Γ is the average channel s/n, viz:

$$\Gamma = \frac{E}{N_o} \tag{24}$$

and N_o is the one-sided spectral density function of the additive white Gaussian noise.

3.2 Weighted PCM System 2

System 1 deploys fixed weighting profiles, where the same profile is used for every PCM word. This is a suboptimum weighting strategy. Observing that as the digital noise power depends on the value of the quantized sample x_i, it appears desirable to have 2^N different weighting profiles, one for each value of i. By this means we are able to select the unique weighting profile that best suits the particular PCM word to be transmitted. Thus, having generated a μ-law PCM word we use the index i to address a Read-Only Memory (ROM) that provides us with the optimum weighting profile $\{\phi_l(x_i)\}_{l=1}^N$ for that particular word. Weighting then ensues and the weighted lth bit has an amplitude

$$B_{2,l} = \sqrt{\phi_l(x_i)}\,b_l; \qquad b_l = \pm 1 \tag{25}$$

for binary amplitude modulation, where the input sample x is quantized to x_i. The PCM decoder for this system, as with the other two weighted PCM systems, is the same as that employed in conventional PCM.

Because the weighting depends upon the value of the quantized level x_i, we amend eq. (23) to

$$\Gamma_l(x_i) = \phi_l(x_i)\Gamma; \qquad \text{System 2.} \tag{26}$$

3.3 Weighted PCM System 3

In System 2 each profile $\{\phi_l(x_i)\}_{l=1}^N$ is optimum for the particular word associated with the quantized level x_i, with the imposed constraint that the word energy for every PCM word is the same. However, the digital noise power varies from word to word, and this leads us to the notion of a subsequent modification to the weighted bit $B_{2,l}$. Specifically, the amplitudes of the weighted bits in the word associated with quantization level x_i are

$$B_{3,l} = \sqrt{W_i}B_{2,l}; \qquad l = 1, 2, \cdots, N, \tag{27}$$

where W_i is a multiplicative factor to the determined. Adopting this approach we arrange for the overall average word energy (and thus the overall average bit energy), rather than the individual word energy, to be constant. We will show that by combining the individual word weighting $\phi_l(x_i)$ of System 2 with the multiplicative factor W_i, a significant enhancement in overall s/n is achieved. In Section 4.3 the constraints on the factor W_i are specified.

The weighting strategy employed in System 3 results in a channel s/n for the lth bit and the ith quantization level of

$$\Gamma_l(x_i) = \phi_l(x_i)W_i\Gamma; \qquad \text{System 3.} \tag{28}$$

3.4 Determining the digital noise power

In order to formulate the digital noise power of eq. (17) we commence with an expression for probability ρ_l that applies for a particular modulation scheme. This probability is dependent upon the long-term channel s/n, Γ, and we replace Γ by one of our values of Γ_l given by either eq. (23), (26), or (28). Thus we obtain an equation of ϵ_a^2 as a function of our weighting factor, namely, ϕ_l, $\phi_l(x_i)$, or $\phi_l(x_i)W_i$. The optimum weighting factor is then found that minimizes the digital noise power, and thereby maximizes the overall s/n. This is the essence of our approach, and in Section IV we present the derivation of the digital noise power in detail.

3.5 *Objective system performance criterion*

The probability of bit error is a poor measure of system performance, as we discussed in Section 4.1. Instead we opt for overall s/n, given by

$$\hat{s/n} = \frac{E\{x^2\}}{\epsilon_q^2 + \epsilon_c^2 + \epsilon_a^2} = \frac{\sigma_x^2}{\epsilon^2}, \tag{29}$$

where $E\{x^2\}$ or σ_x^2 is the mean signal power, and ϵ_q^2, ϵ_c^2, and ϵ_a^2 are the quantization, clipping, and transmission error noise power components of ϵ^2, respectively. The derivation of eq. (29) is given in Ref. 3. By assuming that N is large, e.g., $N = 8$, the noise power generated in the μ-law PCM encoder is $\epsilon_q^2 + \epsilon_c^2$. The quantization noise power ϵ_q^2 is produced in the quantization process when the input variable x is within the range of the quantizer, namely, -1, $+1$. Clipping noise ϵ_c^2 is produced when x exceeds the range of the quantizer as the recovered signal amplitudes are truncated. The channel noise power ϵ_a^2 is added to the encoder noise power to yield the total noise power.

IV. WEIGHTED PCM FOR FADING CHANNELS

The weighting profiles for PCM when the transmission is over channels subjected to Rayleigh fading will now be determined. Unless otherwise stated, it will be assumed that NCFSK modulation is employed.

4.1 *Weighted PCM System 1*

Let us weight the bits in each word using the same profile for a given channel s/n. The bits are scrambled prior to transmission and on descrambling at the receiver the burst errors on the channel manifest as independent random errors. For NCFSK modulation the average bit error probability for M-fold diversity with Ideal Maximal Ratio Combining (IMRC) is[6]

$$P = \frac{2^{M-1}}{(2 + \Gamma)^M}, \tag{30}$$

where Γ is the average per branch channel s/n. From eqs. (23) and (30) we identify ρ_l of eq. (11) as

$$\rho_l = \frac{2^{M-1}}{(2 + \phi_l \Gamma)^M}, \tag{31}$$

and hence from eqs. (16) and (17) the digital noise power becomes

$$\epsilon_a^2 = \sum_{l=1}^{N} \sum_{i=1}^{2^N-1} \frac{2^{M-1}}{(2 + \phi_l \Gamma)^M} p_i a_l(x_i) = \sum_{l=1}^{N} \frac{2^{M-1} A_l}{(2 + \phi_l \Gamma)^M}, \qquad (32)$$

where the constraint of eq. (22) is applicable. In Appendix A, ϵ_a^2 is minimized if the weights

$$\phi_l = \frac{(A_l)^{\frac{1}{M+1}}}{\frac{1}{N} \sum_{k=1}^{N} (A_k)^{\frac{1}{M+1}}} + \frac{2}{\Gamma} \left(\frac{(A_l)^{\frac{1}{M+1}}}{\frac{1}{N} \sum_{k=1}^{N} (A_k)^{\frac{1}{M+1}}} - 1 \right) \qquad (33)$$

are employed. Observe that when no diversity is used M is unity, and for this case the average bit error probability $P = 1/(2 + \Gamma)$ is upper bounded by $P \leq 1/\Gamma$. This bound is tight for channel s/n values of interest, and calculating the weight profile ϕ_l using the upper bound yields the first term in eq. (33), which is independent of Γ. The approximate weighting profile that assumes the upper bound of P for $M = 1$ is also the optimum weights when Γ approaches infinity.

Substituting ϕ_l of eq. (33) into eq. (32) yields the digital noise power

$$\epsilon_a^2 = \frac{2^{M-1}}{(2 + \Gamma)^M} \left(\frac{1}{N}\right)^M \left[\sum_{l=1}^{N} (A_l)^{\frac{1}{M+1}}\right]^{M+1}. \qquad (34)$$

For weights other than those of eq. (33), the above expression is a lower bound on the digital noise power. The average bit error probability from eq. (31) is

$$P_{av} = \frac{1}{N} \sum_{l=1}^{N} \rho_l = \frac{1}{N} \sum_{l=1}^{N} \frac{2^{M-1}}{(2 + \Gamma \phi_l)^M}. \qquad (35)$$

The large contributions to P_{av} derive from the least significant bits because their energy is less than in the unweighted case. The average bit error probability P_{av} for the weighted system does not equal P, and in general it is not a useful design parameter, since the effect of different bit errors on the recovered signal $x_{i,l}$ is radically different. Therefore, we avoid comparing the various weighted PCM systems on the basis of average bit error probabilities. Rather we make our system comparisons at the same average channel s/n, and use as our performance measure the overall s/n given by eq. (29). The $\widehat{\text{s/n}}$ is a good measure of this quality, and although for a given channel s/n the weighting may increase P_{av}, it nearly always reduces ϵ_a^2, and hence increases $\widehat{\text{s/n}}$.

We point out that it is not essential to have N different bit weights to obtain significant reductions in ϵ_a^2 compared to unweighted PCM. Suboptimum weighting schemes with fewer weights than N are discussed in Ref. 2, but we will confine ourselves here to the optimum weighting profile condition.

On rare occasions the weighting profile may include negative weights, for example, when the channel s/n, Γ, is low in eq. (33). The occurrence of negative weights is primarily due to the oversimplified optimization procedures we adopted. Ideally the optimization should always include the constraint $\phi_l \geqslant 0$, $l = 1, 2, \cdots, N$, even when we choose to ignore the effect of double errors, which become more probable for low Γ's and small weights. However, for high-channel signal-to-noise ratios the simplified optimization procedure employed here is quite sufficient. Negative weights can be avoided without this constraint if we elect to discard the second term in eq. (33), i.e., we employ the optimum weights appropriate for high-channel s/n. Of course, the low-channel s/n's that result in a weight becoming negative are of no practical value, as the overall s/n, $\widehat{s/n}$, is then unacceptably low. However, the unrestricted optimization we have performed always provides upper bounds on the gains in $\widehat{s/n}$ that apply to the different weighting schemes; bounds that are reasonable estimates of the gains in $\widehat{s/n}$ that can be practically expected. Thus, in our simulations we employed the asymptotic weight, i.e., the first term in eq. (33), namely,

$$\phi_l = \frac{(A_l)^{\frac{1}{M+1}}}{\frac{1}{N}\sum_{k=1}^{N} (A_k)^{\frac{1}{M+1}}} . \tag{36}$$

This weight was also employed in our simulations using CPSK modulation.

4.2 Weighted PCM System 2

The individual word weighting profiles are dependent upon the $a_l(x_i)$ terms in the A-factors [see eqs. (14) and (16)]. These $a_l(x_i)$ functions enable us to know how the effect of an error in the lth bit varies with the input level x_i. The reader is referred to Ref. 7 for a detailed account of the variation of $a_l(x_i)$ as a function of x_i for $l = 1, 2, \cdots, 8$, for 8-bit μ-law PCM, $\mu = 255$.

We now consider the digital noise power that occurs in this weighting scheme. First, the value of $\Gamma_l(x_i)$ given by eq. (26) replaces Γ in eq. (30). The expression for ρ_l so formulated is then substituted into eq. (17) to give the digital noise power of

$$\epsilon_a^2 = \sum_{i=0}^{2^N-1} p_i \sum_{l=1}^{N} \frac{2^{M-1} a_l(x_i)}{(2 + \phi_l(x_i)\Gamma)^M} . \tag{37}$$

Minimizing ϵ_a^2 with respect to $\phi_l(x_i)$ yields the optimum weights

$$\phi_l(x_i) = \frac{[a_l(x_i)]^{\frac{1}{M+1}}}{\frac{1}{N}\sum_{k=1}^{N}[a_k(x_i)]^{\frac{1}{M+1}}} + \frac{2}{\Gamma}\left[\frac{[a_l(x_i)]^{\frac{1}{M+1}}}{\frac{1}{N}\sum_{k=1}^{N}[a_k(x_i)]^{\frac{1}{M+1}}} - 1\right]. \quad (38)$$

If we use the optimum weights specified by eq. (38), the digital noise power is given by

$$\epsilon_a^2 = \frac{2^{M-1}}{(2+\Gamma)^M}\left(\frac{1}{N}\right)^M \sum_{i=0}^{2^N-1} p_i \left[\sum_{l=1}^{N}[a_l(x_i)]^{\frac{1}{M+1}}\right]^{M+1}. \quad (39)$$

For high-channel s/n only the first term in eq. (38) need be considered. The variation of $\phi_l(x_i)$ as a function of quantized level x_i for 8-bit μ-law PCM, $\mu = 255$, is displayed in Fig. 2 for each bit in the PCM word. The discontinuities in $\phi_l(x_i)$ derive from the discontinuities in $\{a_k(x_i)\}_{k=1}^N$. In the case of $\phi_1(x_i)$, where $a_1(x_i)$ is a monotonic function, the summation of $a_k(x_i)$ in the denominator of $\phi_l(x_i)$ is responsible for the jumps observed in $\phi_1(x_i)$ in Fig. 2a. When diversity is applied, $M = 2$, the word weighting profiles $\phi_l(x_i)$ for System 2 are altered, although they have similar shapes to those shown in Fig. 2.

The average bit error probability is

$$P_{av} = \sum_{i=0}^{2^N-1} p_i \frac{1}{N}\sum_{l=1}^{N}\frac{2^{M-1}}{[2+\phi_l(x_i)\Gamma]^M}. \quad (40)$$

4.3 Weighted PCM System 3

Let us commence our investigation into this form of weighting by introducing the constraint

$$\sum_{i=0}^{2^N-1} p_i \frac{E_i}{N_o} = \frac{E}{N_o}, \quad (41)$$

where p_i is given by eq. (15). We introduce the word weighting factor W_i by ensuring that the channel s/n for x_i is

$$\frac{E_i}{N_o} = W_i\frac{E}{N_o}; \qquad i = 0, 1, 2, \cdots, 2^N - 1. \quad (42)$$

From eqs. (41) and (42) we formulate our imposed constraint on the word weighting factor W_i, viz:

$$\sum_{i=0}^{2^N-1} p_i W_i = 1. \quad (43)$$

The individual bit weights in the PCM words are given by eq. (38), and the previous constraint of eq. (22) with $\phi_l = \phi_l(x_i)$ still applies. We observe that although the word energy is no longer constant, the

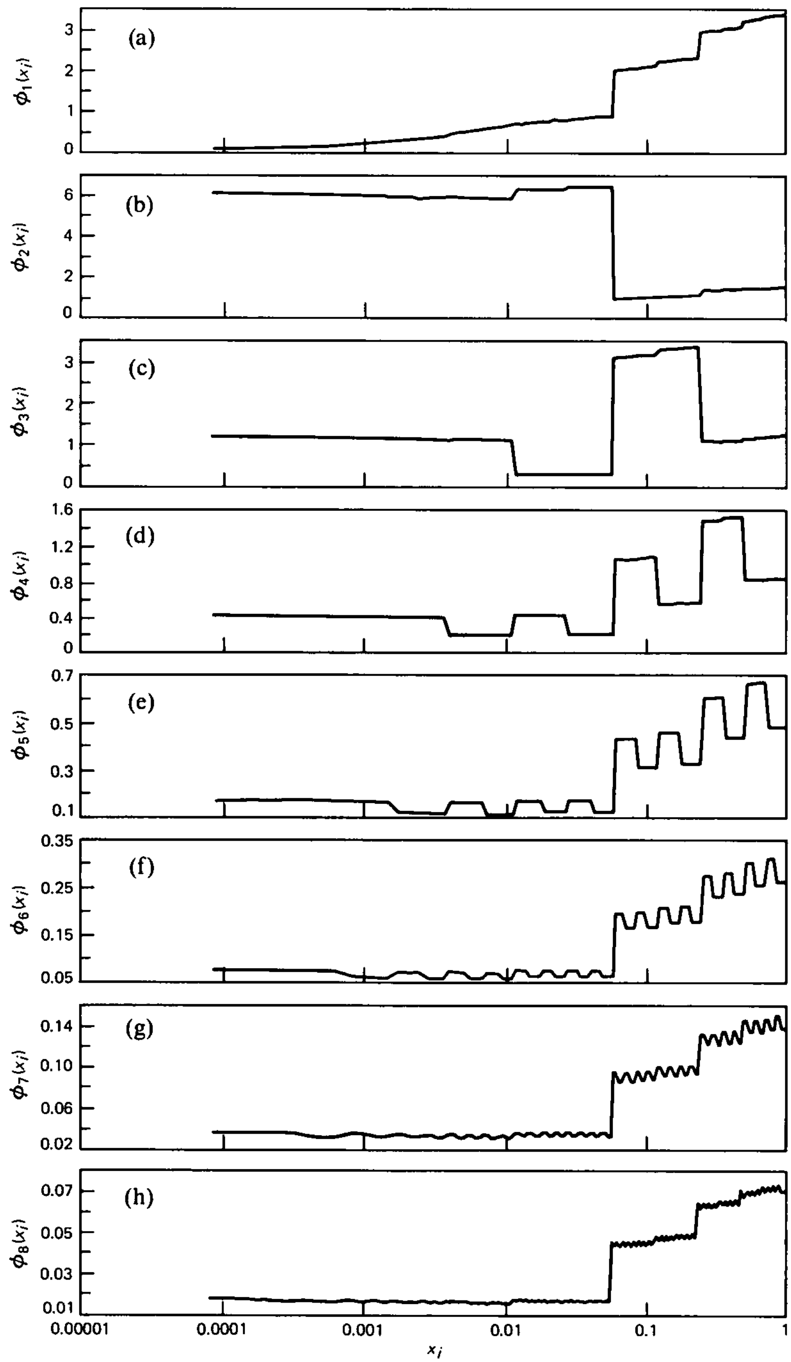


Fig. 2—The variation of the word weighting profiles $\phi_l(x_i)$ as a function of x_i for 8-bit μ-law Rayleigh fading channel PCM, μ = 255, input signal level of −17 dB. The subfigures (a), (b), $\cdots$, (h), relate to l = 1, 2, $\cdots$, 8, respectively.

average PCM word energy E is unchanged by this subsequent word weighting of W_i.

The minimization of the digital noise power is divided into two steps. The optimum bit weight profiles, $\phi_l(x_i)$, are basically the same as in the previous section, with the exception that the word weights are now varied with the "word s/n", i.e., E_i/N_o. The average digital noise power when the weighting set $\{W_i\}$ is included is, from eqs. (17), (28), and (30),

$$\epsilon_a^2 = \sum_{i=0}^{2^N-1} p_i \sum_{l=1}^{N} \frac{2^{M-1} a_l(x_i)}{[2 + W_i \phi_l(x_i)\Gamma]^M} . \tag{44}$$

The optimum weights $\phi_l(x_i)$ given by eq. (38) enable ϵ_a^2 of eq. (44) to be expressed as

$$\epsilon_a^2 = \sum_{i=0}^{2^N-1} p_i \left\{\frac{2^{M-1}}{(2 + W_i\Gamma)^M}\right\} \beta(x_i), \tag{45}$$

where

$$\beta(x_i) = \left(\frac{1}{N}\right)^M \left[\sum_{l=1}^{N} [a_l(x_i)]^{\frac{1}{M+1}}\right]^{M+1} . \tag{46}$$

The function $\beta(x_i)$ depends upon the magnitude of the quantized level and upon $a_l(x_i)$, whose discontinuities cause $\beta(x_i)$ to have sharp jumps as x_i changes segments.

The optimum word weights W_i are determined by minimizing ϵ_a^2 with respect to W_i, employing the constraint of eq. (43). By this method the optimum word weights are

$$W_i = \frac{[\beta(x_i)]^{\frac{1}{M+1}}}{\sum_{k=0}^{2^N-1} p_k[\beta(x_k)]^{\frac{1}{M+1}}} + \frac{2}{\Gamma}\left[\frac{[\beta(x_i)]^{\frac{1}{M+1}}}{\sum_{k=0}^{2^N-1} p_k[\beta(x_k)]^{\frac{1}{M+1}}} - 1\right] . \tag{47}$$

Substituting W_i into eq. (45) yields the digital noise power

$$\epsilon_a^2 = \frac{2^{M-1}}{(2+\Gamma)^M}\left[\sum_{i=0}^{2^N-1} p_i[\beta(x_i)]^{\frac{1}{M+1}}\right]^{M+1} . \tag{48}$$

The average bit error probability is

$$P_{av} = \sum_{i=0}^{2^N-1} p_i \sum_{l=1}^{N} \frac{1}{N} \frac{2^{M-1}}{(1 + \phi_l(x_i) W_i \Gamma)^M} . \tag{49}$$

4.4 Weighting for ideal selection combining diversity and NCFSK

When Ideal Selection Combining (ISC) diversity is employed, the average bit error probability of an unweighted PCM system can be shown to be[6–8]

$$P = \frac{M}{2\Gamma} \sum_{k=0}^{M-1} \binom{M-1}{k} (-1)^k \frac{1}{\left(\frac{1}{\Gamma} + \frac{k}{\Gamma}\right) + \frac{1}{2}}, \tag{50}$$

where M is the order of the diversity. The channel s/n Γ is changed according to the method of weighting, as described in Sections 3.1 through 3.3. For Systems 1, 2, and 3, Γ is replaced by $\phi_l\Gamma$, $\phi_l(x_i)\Gamma$, and $\phi_l(x_i)W_i\Gamma$, respectively. Satisfactory values of the weights are the asymptotic weights derived for the different weighting procedures for ideal maximal ratio combining diversity. Thus, the weights given by eq. (36) are used for System 1; $\phi_l(x_i)$ of eq. (38) is used when Γ tends to infinity [i.e., the asymptotic values of $\phi_l(x_i)$, for System 2], while the same $\phi_l(x_i)$ used for System 2 is also used for System 3, together with the asymptotic value of W_i [see eq. (47)]. As before, the values of ρ_l so determined are substituted into eq. (17) to give the digital noise power for the three systems.

4.5 Weighting for CPSK modulation

When Ideal Maximal Ratio Combining (IMRC) diversity is employed with CPSK, the average bit error probability is[6–8]

$$P = \frac{1}{2}\left[1 - \frac{1}{\sqrt{1 + \frac{1}{\Gamma}}} \sum_{k=1}^{M} \binom{2k-2}{k-1} \left(\frac{1 - \frac{1}{1 + \frac{1}{\Gamma}}}{4}\right)^{k-1}\right] \tag{51}$$

The value of Γ in the above equation is changed according to eq. (23), (26), or (28), depending on whether System 1, 2, or 3 is employed. We employed the asymptotic weights derived for NCFSK and IMRC diversity for this case of CPSK modulation. The probability ρ_l computed by this means is substituted into eq. (17) to yield the digital noise power.

For the case of ISC diversity and CPSK modulation the average bit error probability is[6–8]

$$P = \frac{1}{2} \sum_{k=0}^{M} (-1)^k \binom{M}{k} \frac{1}{\sqrt{1 + \frac{k}{\Gamma}}}. \tag{52}$$

We convert P into ρ_l as before by replacing Γ with the appropriate value given by eq. (23), (26), or (28), and then use the asymptotic weights derived for the case of NCFSK and IMRC diversity.

4.6 Unweighted PCM

For conventional PCM, i.e., where weighting is not employed, and bit scrambling is provided prior to transmission, the modulation being NCFSK, then the digital noise power can be shown to be[7]

$$\epsilon_a^2 = \sum_{w=1}^{N} T_w \left(\frac{2^{M-1}}{(2+\Gamma)^M}\right)^w \tag{53}$$

for IMRC diversity, where

$$T_w = \sum_{j=1}^{w} S_w \binom{N-j}{w-j} (-1)^{w-j}; \qquad w = 1, 2, \cdots, N, \tag{54}$$

and S_w is the sum of the A-factors associated with error sequences containing w ones. When ISC diversity is employed, the digital noise power is[7]

$$\epsilon_a^2 = \sum_{w=1}^{N} T_w \left[\frac{M}{2\Gamma} \sum_{j=0}^{M-1} (-1)^j \binom{M-j}{j} \frac{1}{\left(\frac{1}{\Gamma} + \frac{j}{\Gamma} + \frac{1}{2}\right)}\right]^w. \tag{55}$$

Equations (53) and (55) are included as bench marks against which the digital noise power of weighted PCM can be measured.

V. WEIGHTED PCM FOR THE GAUSSIAN CHANNEL

For this channel there is no need for diversity, nor bit scrambling prior to transmission, as the bit errors are independent in nature.

5.1 Weighted PCM System 1

For NCFSK modulation and unweighted PCM the average bit error probability is

$$P = \frac{1}{2} e^{-\frac{E}{2N_o}}, \tag{56}$$

and when bit weighting is applied such that the lth bit, $l = 1, 2, \cdots, N$, has its amplitude individually adjusted, the expression for P is modified to eq. (23), viz:

$$\rho_l = \frac{1}{2} e^{-\phi_l \frac{E}{2N_o}}. \tag{57}$$

Substituting ρ_l into eq. (17) gives the digital noise power ϵ_a^2. In Appendix B we show that the optimum weights that minimize this digital noise power are

$$\phi_l = 1 + \frac{\ln\left(\dfrac{A_l}{A_o}\right)}{\dfrac{E}{2N_o}}; \qquad l = 1, 2, \cdots, N, \tag{58}$$

where

$$A_o = \left(\prod_{l=1}^{N} A_l\right)^{1/N}. \tag{59}$$

The optimum weights ϕ_l are thus dependent on the channel s/n, namely E/N_o. Substituting the optimum weights ϕ_l into eq. (17) yields

$$\epsilon_a^2 = NA_oP, \tag{60}$$

where P is the average bit error probability given by eq. (56).

The average bit error probability for the unweighted scheme is, from eq. (57),

$$P_{av} = \frac{1}{N}\sum_{l=1}^{N}\frac{1}{2}e^{-\phi_l \frac{E}{2N_o}}. \tag{61}$$

When weighting is applied, and the weighting is optimum at each channel s/n, the average bit error probability is determined by substituting ϕ_l from eq. (58) into eq. (61), viz:

$$P_{av} = \frac{P}{N}\sum_{l=1}^{N}\left(\frac{A_o}{A_l}\right), \tag{62}$$

where P is the average bit error probability for NCFSK when no weighting is used [see eq. (56)].

5.2 Weighted PCM System 2

As the individual word weighting profile ϕ_l depends in this weighting scheme on the value of x_i, we express ϕ_l as $\phi_l(x_i)$, $l = 1, 2, \cdots, N$, for each of the 2^N values of x_i. In Appendix C the weighting profile is shown to be

$$\phi_l(x_i) = 1 + \frac{\ln\left(\dfrac{a_l(x_i)}{a_o(x_i)}\right)}{\dfrac{E}{2N_o}}$$

$$l = 1, 2, \cdots, N,$$

$$i = 0, 1, 2, \cdots, 2^N - 1, \tag{63}$$

where

$$a_o(x_i) = \left(\prod_{l=1}^{N} a_l(x_i) \right)^{1/N} . \tag{64}$$

Thus we have the most significant bit (MSB), next MSB, $\cdots$, LSB, weighted by $\phi_1(x_i)$, $\phi_2(x_i)$, $\cdots$, $\phi_N(x_i)$, respectively, for quantized level x_i, and these word weighting profiles are available for the 2^N values of i, i.e., the 2^N different PCM words. When these optimum weights $\phi_l(x_i)$ are employed, the average digital noise power is, from eqs. (17) and (63),

$$\epsilon_a^2 = P \sum_{i=0}^{2^N-1} p_i N a_o(x_i). \tag{65}$$

The average bit error probability for the optimum individual weighting profiles is, with the aid of eqs. (61) and (63),

$$P_{av} = P \sum_{i=0}^{2^N-1} \frac{p_i}{N} \sum_{l=1}^{N} \frac{a_o(x_i)}{a_l(x_i)} , \tag{66}$$

where P is the bit error probability for NCFSK.

5.3 Weighted PCM System 3

The average digital noise power for System 3 is from eqs. (17), (28), and (56):

$$\epsilon_a^2 = \sum_{i=0}^{2^N-1} p_i N a_o(x_i) \frac{1}{2} e^{-\frac{W_i E}{2N_o}}. \tag{67}$$

The word weighting factors are optimized such that ϵ_a^2 is minimized. This minimization is performed in Appendix D, from which

$$W_i = 1 + \frac{2N_o}{E} \left\{ \ln[Na_o(x_i)] - \sum_{k=0}^{2^N-1} p_k \ln[Na_o(x_k)] \right\}$$

$$i = 0, 1, 2, \cdots, 2^N - 1. \tag{68}$$

The final optimum bit weight is found by multiplying W_i with $\phi_l(x_i)$ of eq. (63), where E is replaced by E_i/W_i. When this is performed the minimum digital noise power becomes

$$\epsilon_a^2 = P e^{\sum_{i=0}^{2^N-1} p_i \ln[Na_o(x_i)]} . \tag{69}$$

The average bit error probability for this combined individual and word-by-word weighting is

$$P_{av} = \sum_{i=0}^{2^N-1} p_i \left\{ \frac{1}{N} \sum_{l=1}^{N} \frac{a_o(x_i)}{a_l(x_i)} \right\} e^{-\frac{W_i E}{2N_o}}. \tag{70}$$

Substituting W_i from eq. (68) into eq. (70) yields

$$P_{av} = P \sum_{i=0}^{2^N-1} \frac{p_i \prod_{k=0}^{2^N-1} [Na_o(x_k)]^{p_k}}{Na_o(x_i)} \cdot \frac{1}{N} \sum_{l=1}^{N} \frac{a_o(x_i)}{a_l(x_i)} . \tag{71}$$

5.4 Unweighted PCM

The digital noise power can be shown to be[7]

$$\epsilon_a^2 = \sum_{w=1}^{N} P^w(1 - P)^{N-w} S_w = \sum_{w=1}^{N} T_w P^w \tag{72}$$

and the value of P depends upon the modulation used. For NCFSK

$$P = \frac{1}{2} e^{-\Gamma/2}, \tag{73}$$

and for CPSK

$$P = Q(\sqrt{2\Gamma}), \tag{74}$$

where $Q(x)$ is the error function

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt. \tag{75}$$

5.5 Peak power constraint

The formulae for the previous weighting schemes have been derived without imposing a limit on the peak power. When such limitations are applied the optimization problem changes, and the digital noise power may increase. Thus our formulae provide upper bounds for the gains in overall s/n, or, alternatively, reductions in digital noise power, compared to unweighted PCM at the same channel s/n.

VI. RESULTS

In this section we present numerical and graphical interpretations of the formulae derived in Section IV. We also provide computer simulation results for speech signals that were subjected to our different weighting strategies and conveyed over simulated Rayleigh fading and Gaussian channels. The encoder was 8-bit μ-law PCM, $\mu = 255$, where the quantized levels were binary encoded using a folded binary code.

6.1 Results for the Rayleigh fading channel

The constant word weighting profile ϕ_l of eq. (36) for $M = 1$, computed using A_l of eq. (10), where $l = 1$ to 8, is presented in Table I for input powers of −17 and −40 dB. The individual word weighting

Table I—Constant word weighting profile $\{\phi_l\}_{l=1}^{N}$ for the Rayleigh fading channel (asymptotic weights)

Input Power	Bit Number l							
	1	2	3	4	5	6	7	8
−17 dB	1.9432	2.5643	2.0376	0.7974	0.3553	0.1732	0.0860	0.0429
−40 dB	0.6156	6.0135	0.7787	0.3250	0.1445	0.0704	0.0349	0.0174

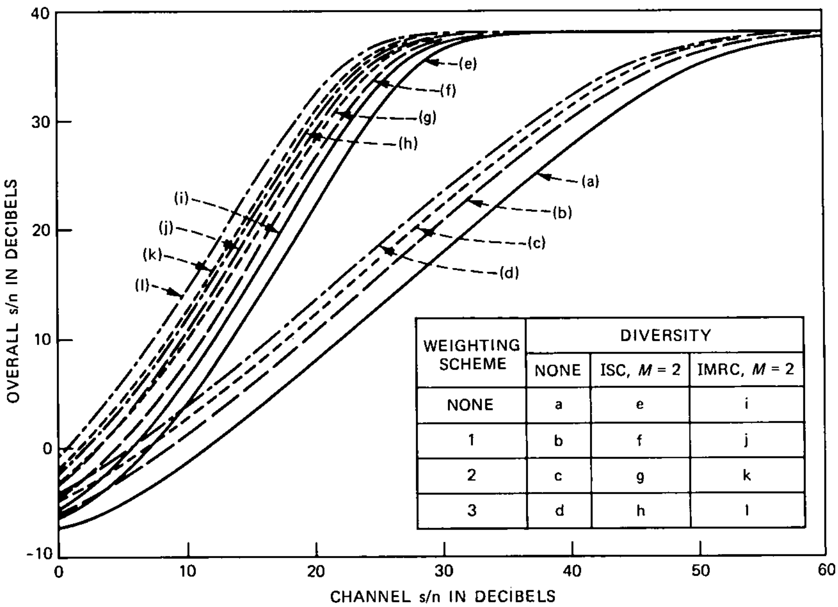


Fig. 3—Rayleigh fading channel. Theoretical curves of overall s/n as a function of channel s/n for NCFSK and an input power level of −17 dB.

profiles $\phi_l(x_i)$ of eq. (38) are displayed in Fig. 2 for $M = 1$. The word weighting factors W_i of eq. (47) are presented in Table II for input powers of −17 and −40 dB, and $M = 1$.

The $\hat{s/n}$ given by eq. (29) is used as our objective performance measure. In the absence of transmission errors ϵ_a^2 is zero, and the $\hat{s/n}$ becomes $\sigma_x^2/(\epsilon_q^2 + \epsilon_c^2)$. For an 8-bit μ-law PCM codec, $\mu = 255$, the variation of $\sigma_x^2/(\epsilon_q^2 + \epsilon_c^2)$ as a function of input power σ_x^2 is determined using well-established results.[2] Two input power levels of −17 and −40 dB relative to the clipping level (±1) are considered, having a correspondence to voiced and unvoiced signal levels, respectively. Knowing σ_x^2 and $\sigma_x^2/(\epsilon_q^2 + \epsilon_c^2)$ enables $(\epsilon_q^2 + \epsilon_c^2)$ to be found. The next task is to compute the digital noise power ϵ_a^2 given by eq. (17). We consider the PDF of the source to be exponential

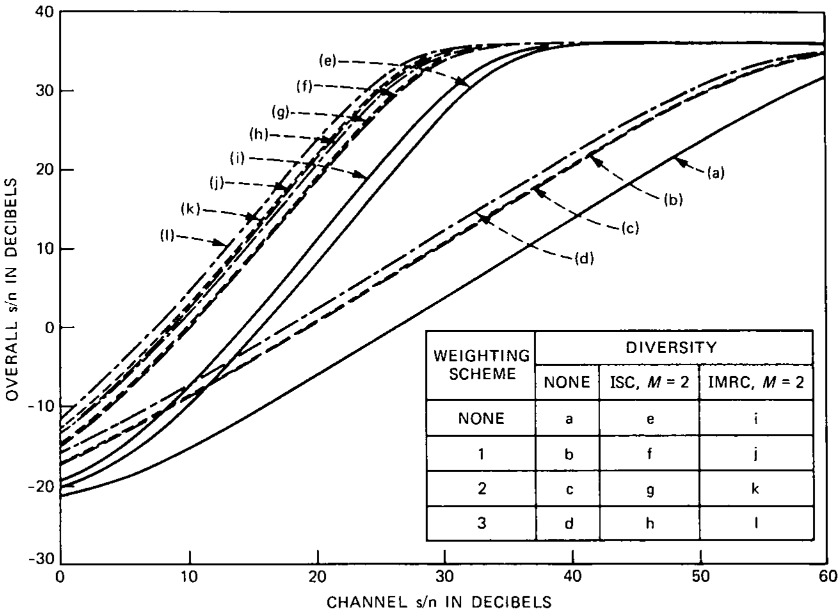


Fig. 4—Rayleigh fading channel. Theoretical curves of overall s/n as a function of channel s/n for NCFSK and an input power level of −40 dB.

$$p_X(x_i) = \frac{1}{\sigma_x \sqrt{2}} \exp\left(-\frac{\sqrt{2} x_i}{\sigma_x}\right), \tag{76}$$

from which the 2^N values of p_i are determined using the 2^N values of x_i. The values of p_i and the $a_l(x_i)$ factors are substituted in eq. (16) to yield the set of A-factors A_l.

The digital noise power ϵ_a^2 of unweighted N-bit μ-law PCM, $N = 8$, $\mu = 255$, is found using either eq. (53) or (55). The variation of ϵ_a^2 as a function of channel s/n is next determined. For a given input power σ_x^2 the $\hat{\text{s/n}}$ depends only on ϵ_a^2, and hence Γ. Consequently, we are able to obtain the graphs of theoretical $\hat{\text{s/n}}$ as a function of channel s/n for unweighted PCM, as shown by curves a in Figs. 3, 4, 5, and 6, where the curves in Figs. 3 and 5 and those in Figs. 4 and 6 apply for input levels of −17 and −40 dB, respectively. These unweighted PCM curves are included as references to which our weighted PCM curves can be compared.

The optimum weights for high-channel s/n values for Systems 1, 2, and 3, i.e., asymptotic weights, are the first term in eqs. (33), (38), and (47), respectively. The digital noise power for the various schemes and conditions applicable to Figs. 3 to 6 is formulated with the aid of eqs. (32), (37), and (44), where the asymptotic weights are used.

Figure 3 shows the curves of theoretical $\hat{\text{s/n}}$ as a function of channel

Table II—Word weight factors W_i for the Rayleigh fading channel (asymptotic weights)

Input Power = −17 dB								
0.1204	0.1258	0.1313	0.1372	0.1432	0.1496	0.1562	0.1632	0.1704
0.1780	0.1859	0.1941	0.2028	0.2118	0.2212	0.2310	0.2410	0.2517
0.2629	0.2746	0.2868	0.2995	0.3128	0.3268	0.3412	0.3564	0.3722
0.3888	0.4060	0.4241	0.4429	0.4626	0.4753	0.4964	0.5185	0.5415
0.5655	0.5907	0.6169	0.6443	0.6727	0.7026	0.7338	0.7664	0.8003
0.8359	0.8729	0.9117	0.9507	0.9929	1.0370	1.0829	1.1310	1.1811
1.2335	1.2882	1.3451	1.4047	1.4670	1.5320	1.5999	1.6708	1.7448
1.8222	0.4804	0.5020	0.5245	0.5481	0.5725	0.5982	0.6250	0.6530
0.6805	0.7110	0.7428	0.7760	0.8104	0.8467	0.8844	0.9239	0.9436
0.9857	1.0297	1.0756	1.1232	1.1733	1.2256	1.2802	1.3337	1.3931
1.4551	1.5199	1.5870	1.6576	1.7313	1.8083	1.2848	1.3422	1.4020
1.4646	1.5289	1.5971	1.6682	1.7426	1.8100	1.8906	1.9747	2.0626
2.1530	2.2488	2.3487	2.4532	2.4303	2.5385	2.6511	2.7690	2.8900
3.0184	3.1523	3.2924	3.4170	3.5688	3.7270	3.8925	4.0622	4.2426
4.4305	4.6272							

Input Power = −40 dB							
0.3664	0.3827	0.3996	0.4174	0.4359	0.4552	0.4754	0.4966
0.5185	0.5416	0.5657	0.5908	0.6171	0.6445	0.6731	0.7031
0.7333	0.7659	0.8000	0.8356	0.8727	0.9115	0.9520	0.9943
1.0384	1.0845	1.1327	1.1830	1.2356	1.2905	1.3478	1.4077
1.4464	1.5107	1.5778	1.6478	1.7210	1.7974	1.8772	1.9605
2.0472	2.1381	2.2329	2.3320	2.4355	2.5435	2.6564	2.7742
2.8932	3.0215	3.1555	3.2954	3.4415	3.5941	3.7535	3.9199
4.0931	4.2745	4.4640	4.6619	4.8685	5.0843	5.3096	5.5449
1.4618	1.5275	1.5961	1.6678	1.7421	1.8202	1.9018	1.9870
2.0708	2.1635	2.2603	2.3613	2.4662	2.5764	2.6914	2.8116
2.8715	2.9996	3.1334	3.2732	3.4181	3.5705	3.7295	3.8956
4.0586	4.2393	4.4279	4.6250	4.8293	5.0442	5.2683	5.5026
3.9097	4.0844	4.2664	4.4568	4.6526	4.8601	5.0764	5.3027
5.5078	5.7532	6.0090	6.2766	6.5515	6.8431	7.1471	7.4560
7.3956	7.7246	8.0675	8.4263	8.7943	9.1852	9.5926	10.0189
10.3980	10.8599	11.3413	11.8449	12.3614	12.9102	13.4822	14.0806

s/n for NCFSK; an input power level of −17 dB; no diversity, and two-branch diversity, when IMRC and ISC techniques are used; and the effect of employing the three weighting schemes. Curves b, c, and d apply to weighting PCM Systems 1, 2, and 3, respectively, when diversity reception is not used. Comparing these curves to the unweighted PCM curve a, we observe that over the range of channel s/n where these are parallel, the gain in $\hat{s/n}$ is 2.8, 4.5, and 6 dB, respectively. When second-order diversity is employed, curves e to l apply. Curves e, f, g, and h relate to unweighted PCM and weighted PCM Systems 1, 2, and 3, respectively; the diversity is ISC. We observe that the gains in $\hat{s/n}$ for Systems 1, 2, and 3 compared to unweighted PCM are significantly larger than those for the case of $M = 1$. Further, the systems are able to operate at a much lower channel s/n, Γ, for the same $\hat{s/n}$. When we use IMRC diversity the gains in $\hat{s/n}$ are similar to those achieved with ISC, but for a given Γ, the values of $\hat{s/n}$ are increased by approximately 2 dB.

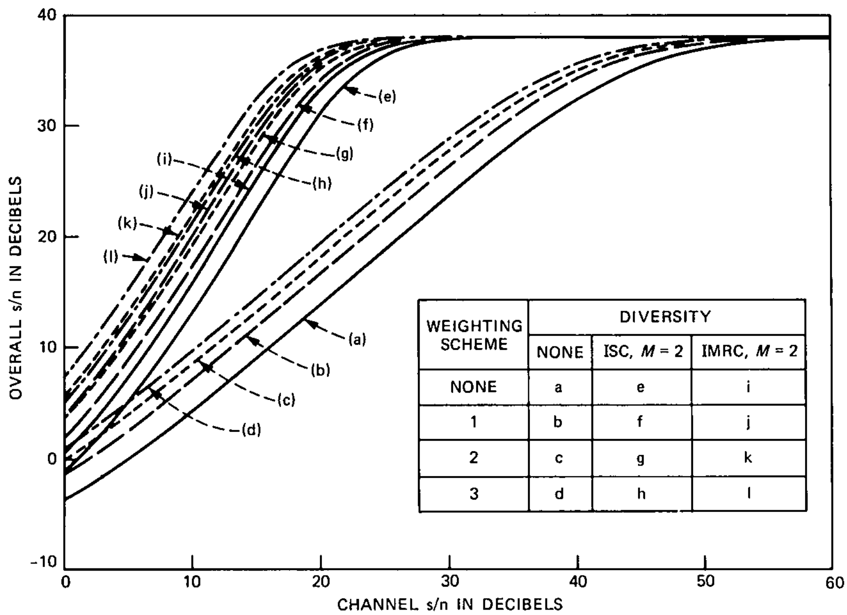


Fig. 5—Rayleigh fading channel. Theoretical curves of overall s/n as a function of channel s/n for CPSK and an input power level of −17 dB.

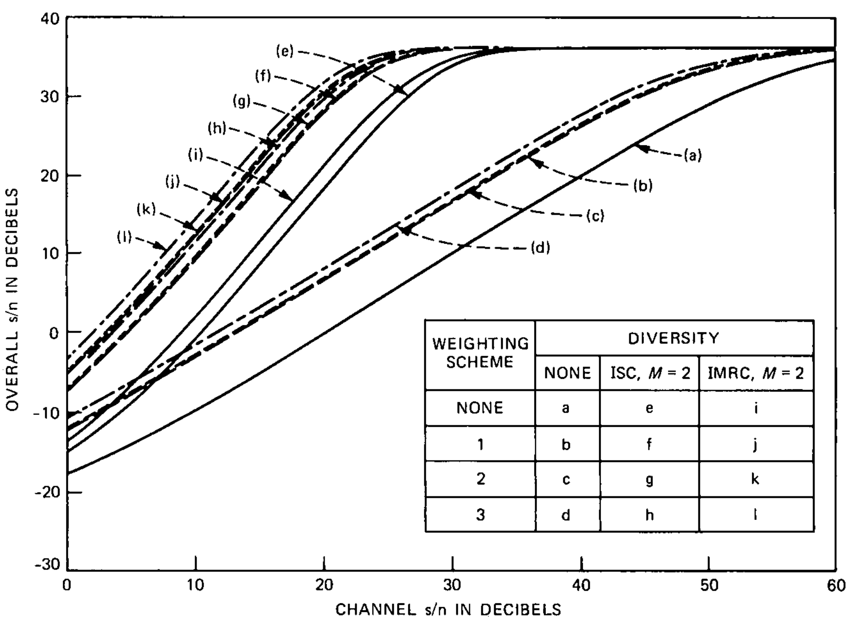


Fig. 6—Rayleigh fading channel. Theoretical curves of overall s/n as a function of channel s/n for CPSK and an input power level of −40 dB.

When the level of the input signal is reduced to −40 dB the curves in Fig. 4 are obtained. Curves a to l relate to the same system parameters as those in Fig. 3. As we expected the peak $\widehat{s/n}$ (at high values of Γ) are reduced by 2 dB due to coarse quantization at low signal levels. The slopes of the curves in Fig. 4 are lower than those in Fig. 3. However, the curves of $\widehat{s/n}$ for Systems 2 and 3 are approximately the same when no diversity is applied. Marginally greater gains in $\widehat{s/n}$ occur at this lower power level of −40 dB when second-order diversity reception is employed.

The theoretical curves for CPSK are shown in Figs. 5 and 6. The curves a to l in Figs. 5 and 6 correspond to those in Figs. 3 and 4, respectively. The essential difference between the curves is that for the same $\widehat{s/n}$, the Γ is 6 dB lower when CPSK is employed compared to when NCFSK modulation is used.

6.1.1 Speech input signals

Speech composed of four concatenated sentences, "Glue the sheet to the dark blue background," "Rice is often served in round bowls," "Four hours of steady work faced us," and "The box was thrown beside the parked truck," were used as our speech input signal. The former two sentences were spoken by females, the latter two sentences by males. The speech signal was bandlimited between 200 to 3200 Hz, sampled at 8 kHz, and stored in the computer. Eight-bit μ-law PCM encoding ensued, the bit stream suitably weighted, and two-level CPSK modulation of a Radio Frequency (RF) carrier performed. From a hardware simulator of frequency-selective Rayleigh-fading mobile radio paths[9] samples of the envelope function $C(t)$ were taken at a rate of 32 kHz and stored in the computer to provide the fading envelope of a mobile unit traveling at 15 mph. By resampling $C(t)$ different vehicle speeds could be simulated. The regenerated lth bit at the kth instant was[10]

$$\begin{aligned} \hat{B}_{(\cdot),l}(k) &= C(k)B_{(\cdot),l}(k) + I(k) \geq 0; \qquad \text{logical 1 generated} \\ \hat{B}_{(\cdot),l}(k) &= C(k)B_{(\cdot),l}(k) + I(k) < 0; \qquad \text{logical 0 generated} \\ l &= 1, 2, \cdots, N, \end{aligned} \tag{77}$$

where $(\cdot)$ is 0, 1, 2, or 3, depending on whether the PCM was unweighted or System 1, 2, or 3 was used, respectively. The amplitude of the Rayleigh envelope, the transmitted bit, and the additive interference level are represented in eq. (77) by $C(k)$, $B_{(\cdot),l}(k)$ and $I(k)$, respectively. After bit regeneration the μ-law PCM words were formulated and subsequently decoded. This decoded speech sequence was subtracted from its counterpart in the input speech sequence to yield the noise sequence. The $\widehat{s/n}$ was computed as the ratio of the input

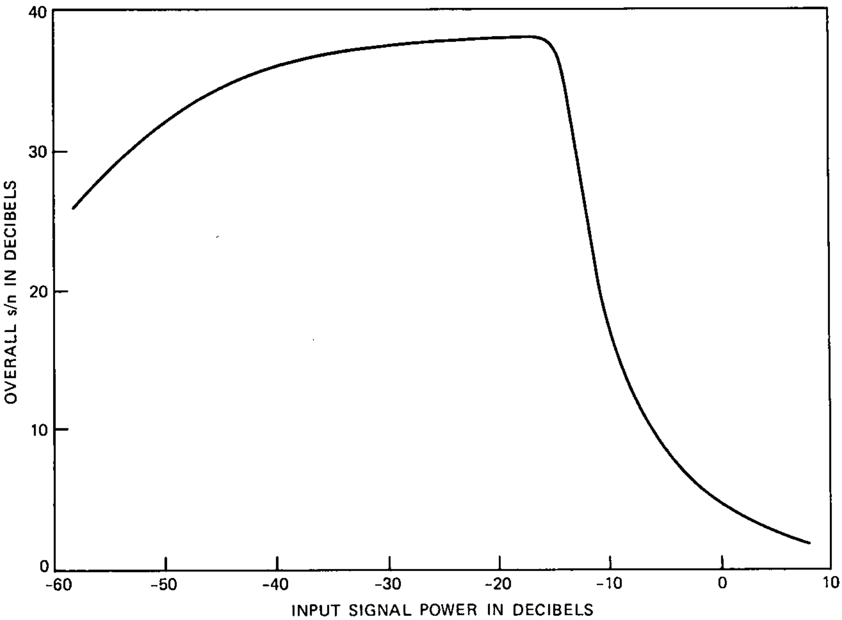


Fig. 7—Overall s/n as a function of input speech power for zero bit error rate.

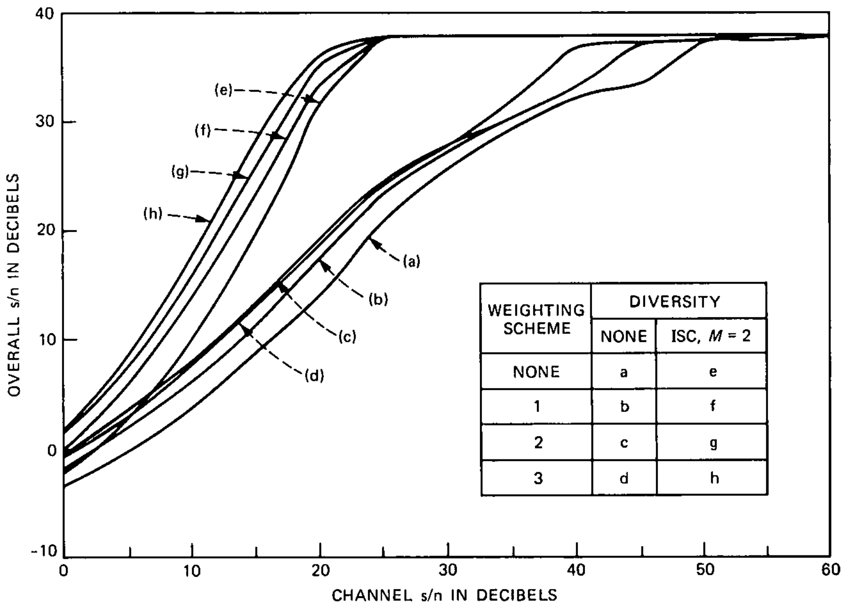
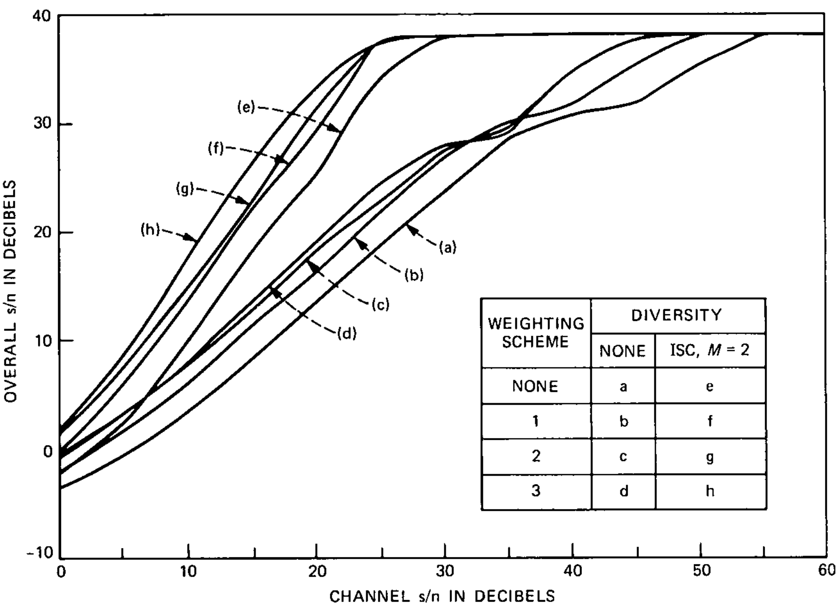


Fig. 8—Mobile radio channel. Overall s/n versus channel s/n for CPSK and speech input signals. The input power level is −17 dB and the vehicular speed is 30 mph.

speech power to the noise sequence power. The curve of $s/\hat{n}$ as a function of input speech power is displayed in Fig. 7 for zero BER.

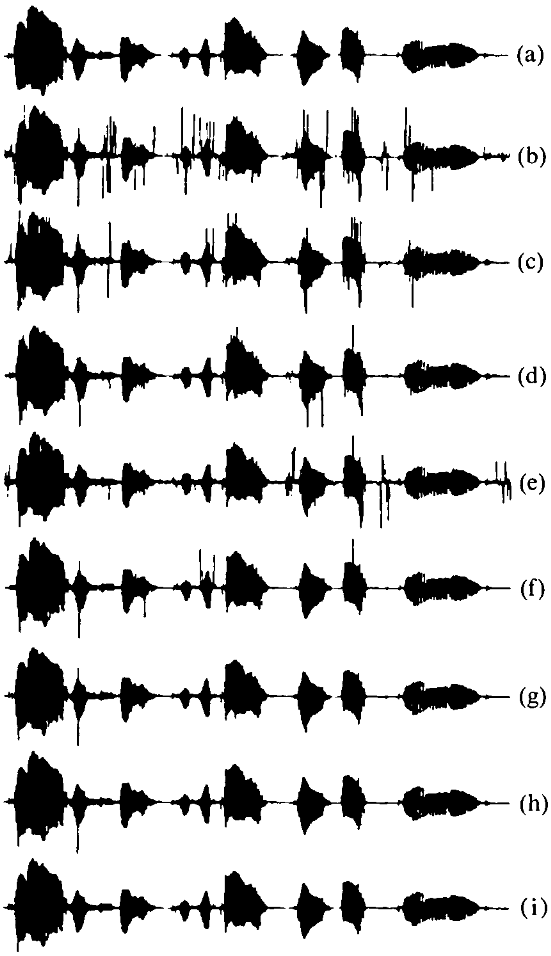
Figure 8 shows the performance of our three weighting systems and unweighted PCM in the absence of diversity and when second-order ISC diversity was employed for an input level of −17 dB. The corresponding performance when the vehicular speech was 60 mph is displayed in Fig. 9. The effects of the nonstationary speech statistics and the duration of the speech and Rayleigh envelope data conspire to cause the curves in Figs. 8 and 9 to lose the parallelism of those in Fig. 5. However, the $s/\hat{n}$ advantage due to using diversity and our weighting systems when speech input signals are employed has a close correspondence with the gains predicted from our theoretical results. We note that the performance of the system as the vehicular speed doubles is, in general, very similar, although there are some large local variations.

Figure 10a shows a segment of speech, while Fig. 10b, c, d, and e were the recovered speech waveforms for unweighted PCM, weighted PCM Systems 1, 2, and 3, respectively, when the vehicular speech was 30 mph, an input level of −17 dB, CPSK modulation, no diversity, $\Gamma = 16$ dB. When ISC, $M = 2$ was employed and the other conditions remained unchanged, the results were the corresponding recovered speech waveforms in Fig. 10 f, g, h, and i. We observe that the effect



WEIGHTING SCHEME	DIVERSITY	
	NONE	ISC, $M = 2$
NONE	a	e
1	b	f
2	c	g
3	d	h

Fig. 9—Mobile radio channel. Overall s/n versus channel s/n for CPSK and speech input signals. The input power level is −17 dB and the vehicular speed is 60 mph.



WEIGHTING SCHEME	DIVERSITY	
	NONE	ISC, $M = 2$
NONE	a	e
1	b	f
2	c	g
3	d	h

Fig. 10—Segment of the speech signal. Overall s/n versus channel s/n for CPSK and speech input signals. The input power level is −17 dB and the vehicular speed is 30 mph. The channel s/n is 16 dB and the original speech segment is waveform i.

of employing Systems 3 and 2 results in a considerable reduction in digital noise power compared to the conventional unweighted PCM.

6.1.2 Average error probability

We assert in Section 4.1 that for weighted PCM systems the average bit error probability P_{av} is not a suitable performance parameter. To lend credence to this statement we display in Fig. 11 the curves of error probability for the MSB, next MSB, $\cdots$, LSB, namely ρ_1, ρ_2, $\cdots$, ρ_8, respectively, as a function of channel s/n, for Weighted PCM System 1. Also shown is the variation of the probability of bit error P for a conventional, unweighted PCM system. The results are for a

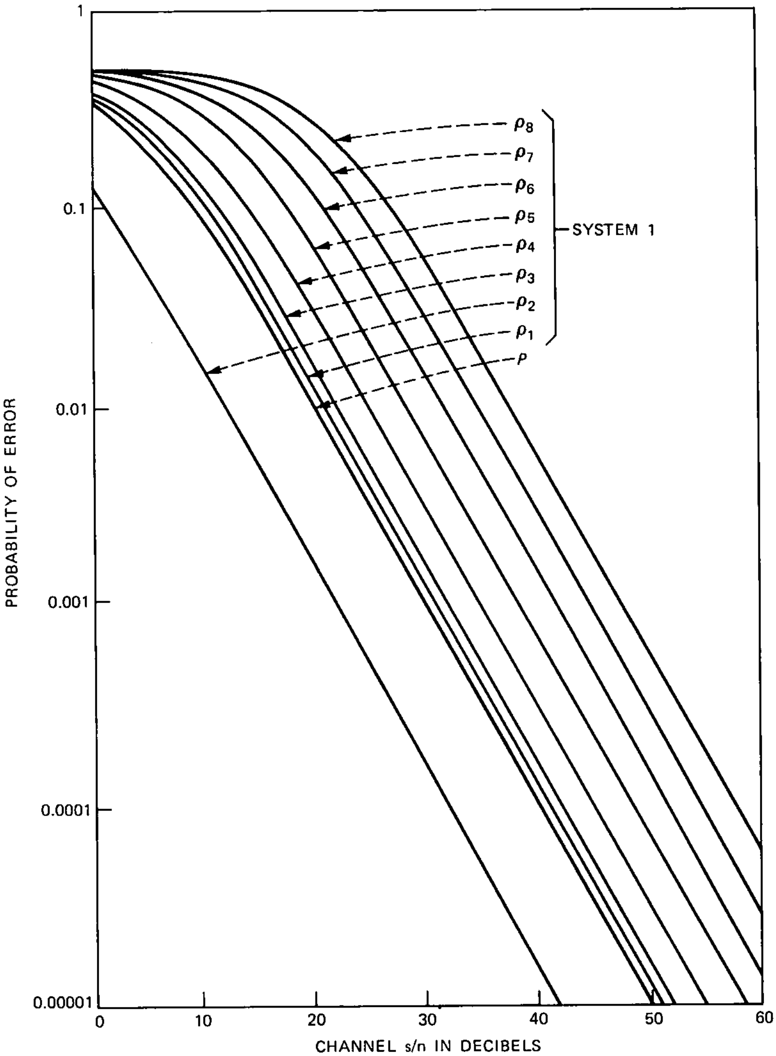


Fig. 11—Curves of probability of error as a functin of channel s/n for the individual probability of bit errors $\rho_1, \rho_2, \cdots, \rho_8$ for System 1. It is a Rayleigh fading channel with no diversity, and in input level of −40 dB. Also displayed is the probability of error P for the unweighted PCM system as a function of channel s/n.

Rayleigh fading channel, no diversity, and an input level of −40 dB. The weighting strategy for System 1 arranges for the LSB to have a lower transmitted energy than the other bits, and, consequently, a greater probability of error. However, an error in the LSB causes the smallest contribution to the digital noise power compared to the contributions from the other bits. We observe that $\rho_8 > \rho_7 > \rho_6 > \rho_5$

$> \rho_4 > \rho_3 > \rho_2$, i.e., the larger the contribution to the digital noise power caused by the regeneration of an erroneous bit, the lower is its probability of being in error. Probability ρ_2 is associated with the most significant magnitude bit of the quantized sample, and we see that $\rho_2 < \rho_1$. This inequality originates in the optimization process, indicating that an error in the most significant magnitude bit will, in general, generate more digital noise than an error in the polarity bit of the reconstituted sample. The probability of error P for the unweighted PCM is seen to be marginally smaller than ρ_1, but significantly greater than ρ_2.

When the average probability P_{av} is computed according to eq. (35), we have the curve displayed in Fig. 12. As we expected from the curves in Fig. 11, P_{av} for Weighted PCM System 1 is considerably in excess of the probability P for unweighted PCM. Also shown in Fig. 12 are P_{av} for Weighted PCM Systems 2 and 3. From our previous figures displaying $s/\hat{n}$, and those depicting P_{av} in Figs. 11 and 12, we conclude that systems that provide superior $s/\hat{n}$ also have inferior P_{av}. Hence we reject P_{av} as a meaningful performance parameter. However, knowledge of P_{av} does contribute to our insight into weighted PCM systems.

6.2 Results for the Gaussian channel

When the fixed weighting profile described in Section 5.1 is used, the digital noise power is given by eq. (60). The optimum weights ϕ_l of eq. (58), having the same A_l factors as used in the unweighted case, are tabulated in Table III. However, we observe that ϕ_l is a function of the channel s/n. Applying the same procedure as for the theoretical results in Section 6.1, we change the weighting ϕ_l with channel s/n to ensure that ϵ_a^2 is minimized. The curves in Figs. 13 and 14 are therefore optimum for every channel s/n. The improvement attained in $s/\hat{n}$ for Weighted PCM Systems 1, 2, and 3 compared to unweighted PCM are 10, 12, and 17 dB, respectively, for an input signal level of −17 dB, respectively, when the channel s/n is 10 dB. The optimum fixed word weighting profile of System 1 enables the channel s/n to deteriorate by 1.5 and 2 dB for input signal levels of −17 and −40 dB, respectively, compared to unweighted PCM, while $s/\hat{n}$ is maintained at 30 dB. Observe that at low values of channel s/n where the BER is high, the application of fixed word weighting can increase $s/\hat{n}$ by 10 to 20 dB. These large gains can transform unintelligible speech into distorted but partially intelligible speech. We note that the gain in $s/\hat{n}$ when the individual word weighting profile is used instead of the fixed weighting profile is only 1 to 2 dB.

The digital noise power ϵ_a^2 for the individual word weighting profiles of System 2 is given by eq. (65), where the $a_l(x_i)$ factors are given in Ref. 7. The individual word weighting profiles of eq. (63) are presented

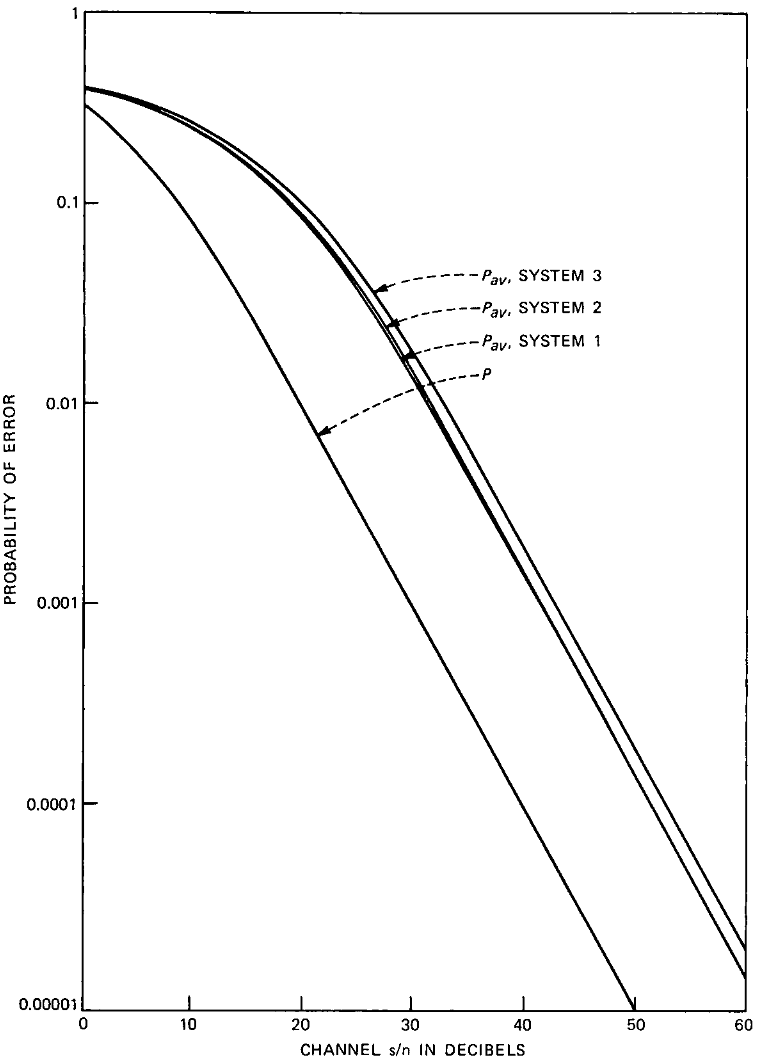


Fig. 12—Curves of average probability for Systems 1, 2, and 3, and the probability of error P for unweighted PCM, as a function of channel s/n.

Table III—The constant word weighting profile $\{\phi_l\}_{l=1}^{N}$ for the Gaussian channel (optimized for a channel s/n of 11 dB)

Input Power	Bit Number l							
	1	2	3	4	5	6	7	8
−17dB	1.4612	1.5493	1.4763	1.1782	0.9213	0.6930	0.4707	0.2500
−40 dB	1.3247	2.0489	1.3994	1.1217	0.8643	0.6355	0.4131	0.1924

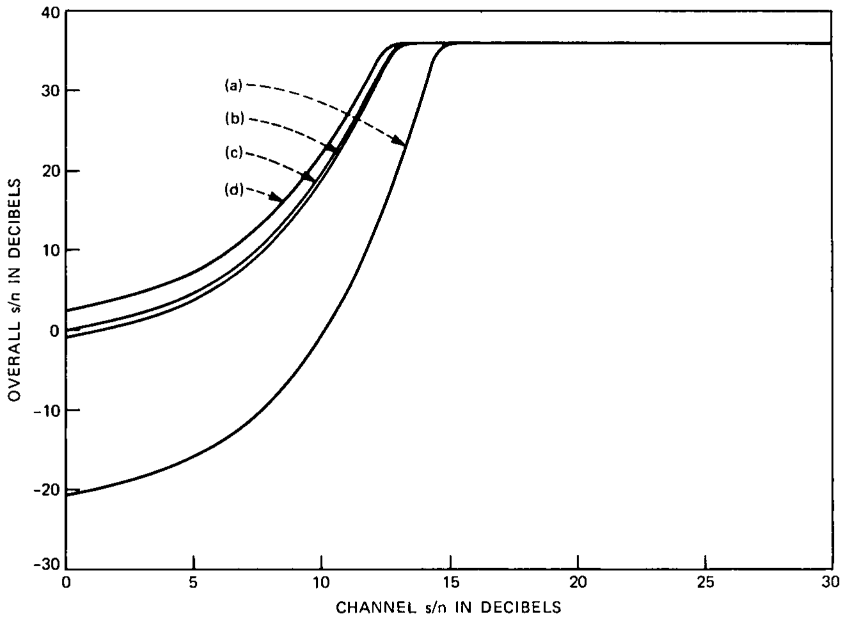


Fig. 13—Gaussian channel. Theoretical curves of overall s/n versus channel s/n for an input level of −17 dB. Curves a, b, c, and d relate to no weighting, weighting scheme 1, 2, and 3, respectively. No diversity was employed.

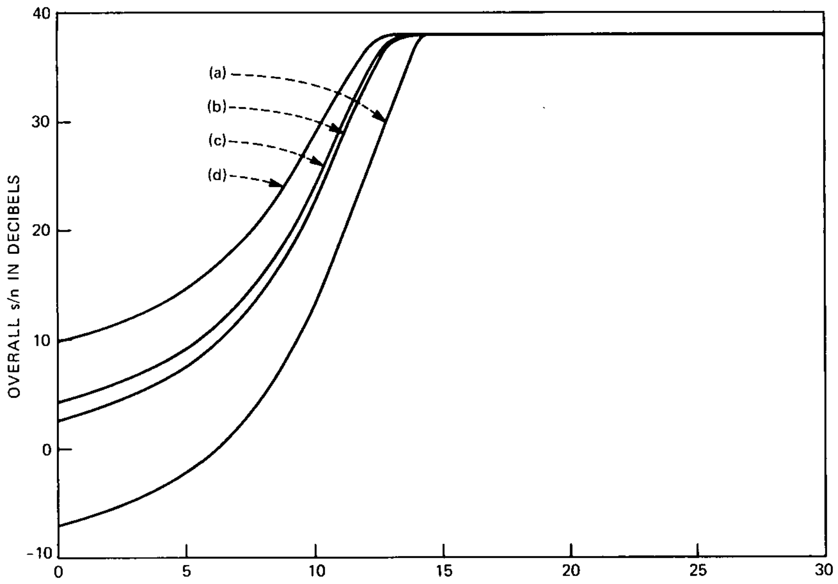


Fig. 14—Gaussian channel. Theoretical curves of overall s/n versus channel s/n for an input level of −40 dB. Curves a, b, c, and d relate to no weighting, weighting scheme 1, 2, and 3, respectively. No diversity was employed.

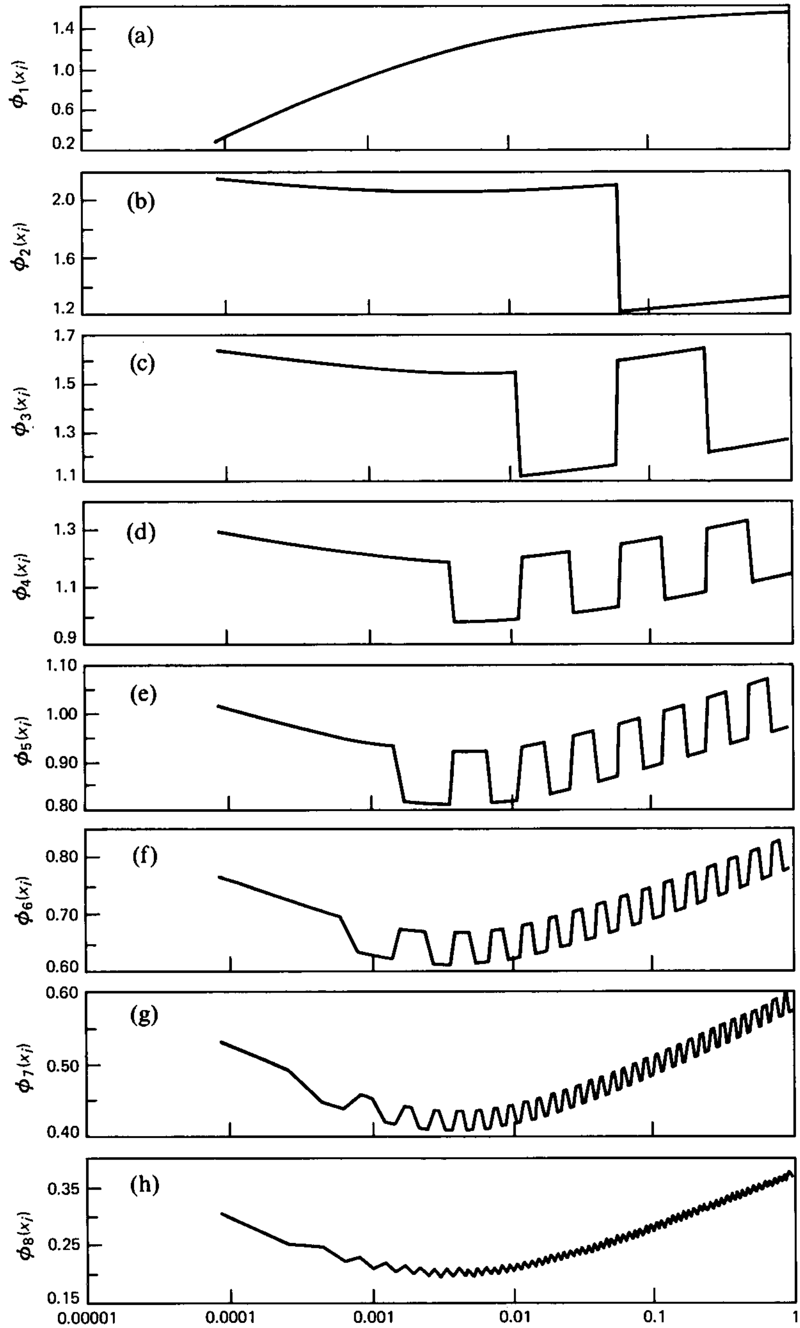


Fig. 15—Gaussian channel. The weighting profiles of Weighting System 2. Subfigures a, b, ⋯, h, corresponding to a single bit error in the polarity bit, the most significant magnitude bit, ⋯, the least significant magnitude bit, respectively.

Table IV—The word weighting factors W_i for the Gaussian channel (optimized for channel s/n of 11 dB)

Input Power = −17 dB								
0.0785	0.1333	0.1648	0.1893	0.2105	0.2297	0.2476	0.2645	0.2807
0.2963	0.3116	0.3264	0.3410	0.3553	0.3694	0.3833	0.3971	0.4107
0.4242	0.4376	0.4509	0.4641	0.4772	0.4903	0.5033	0.5162	0.5291
0.5419	0.5547	0.5674	0.5801	0.5927	0.6054	0.6179	0.6305	0.6430
0.6555	0.6680	0.6805	0.6929	0.7053	0.7177	0.7301	0.7424	0.7548
0.7671	0.7794	0.7917	0.8040	0.8163	0.8286	0.8408	0.8531	0.8653
0.8775	0.8898	0.9020	0.9142	0.9264	0.9385	0.9507	0.9629	0.9751
0.9872	0.9994	1.0116	1.0237	1.0358	1.0480	1.0601	1.0723	1.0844
1.0965	1.1086	1.1207	1.1329	1.1450	1.1571	1.1692	1.1813	1.1934
1.2055	1.2176	1.2297	1.2418	1.2538	1.2659	1.2780	1.2901	1.3022
1.3143	1.3263	1.3384	1.3505	1.3626	1.3746	1.3867	1.3988	1.4108
1.4229	1.3450	1.4470	1.4591	1.4712	1.4832	1.4953	1.5074	1.5194
1.5315	1.5435	1.5556	1.5677	1.5797	1.5918	1.6038	1.6159	1.6279
1.6400	1.6521	1.6641	1.6762	1.6882	1.7003	1.7123	1.7244	1.7364
1.7485	1.7605							
Input Power = −40 dB								
0.6457	0.7005	0.7320	0.7566	0.7778	0.7970	0.8148	0.8317	0.8479
0.8636	0.8788	0.8937	0.9082	0.9226	0.9367	0.9506	0.9644	0.9870
0.9915	1.0049	1.0182	1.0314	1.0445	1.0576	1.0705	1.0835	1.0963
1.1092	1.1219	1.1347	1.1473	1.1600	1.1726	1.1852	1.1978	1.2103
1.2228	1.2353	1.2477	1.2602	1.2726	1.2850	1.2973	1.3097	1.3220
1.3344	1.3467	1.3590	1.3713	1.3836	1.3958	1.4081	1.4203	1.4326
1.4448	1.4570	1.4692	1.4814	1.4936	1.5058	1.5180	1.5302	1.5423
1.5545	1.5667	1.5788	1.5910	1.6031	1.6152	1.6274	1.6395	1.6516
1.6638	1.6759	1.6880	1.7001	1.7122	1.7243	1.7364	1.7485	1.7606
1.7727	1.7848	1.7969	1.8090	1.8211	1.8332	1.8453	1.8574	1.8694
1.8815	1.8936	1.9057	1.9178	1.9298	1.9419	1.9540	1.9660	1.9781
1.9902	2.0022	2.0143	2.0264	2.0384	2.0505	2.0626	2.0746	2.0867
2.0987	2.1108	2.1229	2.1349	2.1470	2.1590	2.1711	2.1832	2.1952
2.2073	2.2193	2.2314	2.2434	2.2555	2.2675	2.2796	2.2916	2.3037
2.3157	2.3278							

in Fig. 15. We used these weighting profiles to generate the curves for Weighting System 2 in Figs. 13 and 14.

When the scheme of Section 5.3 is used the $\widehat{s/n}$ is computed using eqs. (69) and (29). The word weighting factors are listed in Table IV. The gain in $\widehat{s/n}$ when the individual word weighting factor W_i is introduced is substantial, as shown by the performance curves of Weighting System 3. These gains in Figs. 13 and 14 exceed 20 dB compared to unweighted PCM when the input signal power is −40 dB and the channel s/n falls below 12 dB.

VII. DISCUSSION

In another paper[7] we described the effect of transmission errors on μ-law PCM encoded signals transmitted over Gaussian and Rayleigh fading channels. The essential theoretical results of that work are condensed here into Sections 4.6 and 5.4, and are employed as bench marks for measuring the gains in overall s/n when the μ-law PCM bits are weighted prior to transmission. Weighting System 1 is not novel. It is the original scheme due to Bedrosian[1] that has been analyzed by him and Sundberg.[2] However, the results for transmission of weighted μ-law PCM by System 1 over Rayleigh fading channels are new. Systems 2 and 3 mark a basic departure from previous proposals for weighting. In System 2 we introduce the concept of instantaneous weighting, where the bits are weighted by a profile dependent on the quantization level. Thus there are 2^N different weighting profiles, where N is the number of bits in the PCM words and these weighting profiles can be conveniently stored in a ROM. Observing that the digital noise power is dependent on the actual PCM words generated leads us to the notion of assembling the bits weighted by System 2 into their binary words and then weighting each bit in the word by a value unique to the word.

We have presented theoretical results of overall s/n as a function of channel s/n for all our weighting schemes when the transmission is over Rayleigh and Gaussian channels, the modulation NCFSK and CPSK, and when ISC and IMRC diversity is employed. The overall s/n, i.e., the $\hat{s/n}$ of the recovered signal, is a meaningful objective measure of system performance. By contrast BER is a misleading performance parameter in weighted PCM. Indeed, the overall s/n generally increases when the BER increases. This is because the large contributions to the average bit error probability arise from the least significant bits that contribute least to the digital noise power.

In addition to the provision of theoretical equations and graphs of overall s/n, we have provided simulation results when the input signal was composed of four concatenated speech sentences. The simulations were limited to two-level CPSK modulation and apply to a mobile radio channel. However, the simulations encompassed all the various weighting schemes, unweighted μ-law PCM, no diversity, and ISC, $M = 2$, diversity. The results for speech agree satisfactorily with our theory, and the waveform presentation of Fig. 10 shows that the enhancement of the recovered waveforms has a matching correspondence with the measured gains in $\hat{s/n}$.

In summary we make the following observations. For the Rayleigh fading channel, and in the absence of diversity, Weighting Systems 1, 2, and 3 yielded gains of $\hat{s/n}$ over unweighted μ-law PCM of approximated 3, 4.5, and 6 dB, when the input power was −17 dB. This was irrespective of whether the modulation was NCFSK or CPSK. At the lower input level of −40 dB, Systems 1 and 2 had the same performance with a gain in $\hat{s/n}$ of 6.5 dB. System 3 enhanced the performance by approximately 8 dB. Second-order diversity significantly extends the range of channel s/n before $\hat{s/n}$ begins to deteriorate. This is, of course, well known, as is the ability of IMRC diversity to outperform ISC in terms of $\hat{s/n}$ for a given channel s/n. What was not known was the behavior of these diversity schemes when weighted PCM was introduced. We found that the relative performance of the three weighting systems with no diversity was repeated with diversity, enabling us to draw the general conclusion that System 3 always performs best, and that at low power levels there is no advantage to using System 2 compared to System 1.

For Gaussian channel the improvement in $\hat{s/n}$ for Systems 1, 2, and 3 compared to unweighted μ-law PCM was 10, 12, and 17 dB, respectively, for a channel s/n of 10 dB, and an input level of −17 dB. Thus a $\hat{s/n}$ for unweighted μ-law PCM of 13 dB was transformed to 30 dB, a value often associated with toll quality $\hat{s/n}$. At the lower input level of −40 dB the gains for Systems 1, 2, and 3, become 20, 21, and 24 dB, respectively, i.e., Systems 1 and 2 had virtually the same performance, while System 3 only yielded a gain of 3 dB over System 2.

REFERENCES

1. E. Bedrosian. "Weighted PCM," IRE Trans. Information Theory, *IT-4* (March 1958), pp. 45–9.
2. C.-E. Sundberg, "Optimum Weighted PCM for Speech Signals," IEEE Trans. Commun., *COM-26,* No. 6 (June 1978), pp. 872–81.
3. N. Rydbeck and C.-E. Sundberg, "Analysis of Digital Errors in Non-linear PCM Systems," IEEE Trans. Commun., *COM-24,* No. 1 (January 1976), pp. 59–65.
4. C.-E. Sundberg and N. Rydbeck, "Pulse Code Modulation with Error-Correcting Codes for TDMA Satellite Communication Systems," Ericsson Technics, *32,* No. 1 (1976) pp. 3–56.
5. C.-E. Sundberg, "The Effect of Single Bit Errors in Standard Nonlinear PCM Systems," IEEE Trans. Commun., *COM-24,* No. 9 (September 1976), pp. 1062–4.
6. W. C. Jakes, Jr., *Microwave Mobile Communications,* New York: Wiley, 1974.
7. R. Steele, C.-E. Sundberg, and W. C. Wong, "Transmission Errors in Companded PCM Over Gaussian and Rayleigh Fading Channels," AT&T Bell Lab. Tech. J., *63,* No. 6 (July–August 1984).
8. M. Schwartz, W. R. Bennett and S. Stein, *Communication Systems and Techniques,* New York: McGraw Hill, 1966.
9. H. W. Arnold and W. F. Bodtmann, "A Hybrid Multichannel Hardware Simulator for Frequency-Selective Mobile Radio Paths," IEEE Trans. Commun., *COM-31,* No. 3 (March 1983), pp. 370–7.
10. W. C. Wong, R. Steele, B. Glance, and D. Horn. "Time Diversity With Adaptive Error Detection to Combat Rayleigh Fading in Digital Mobile Radio," IEEE Trans. Commun., *COM-31,* No. 3 (March 1983), pp. 378–87.
11. J. M. Chambers, *Computational Methods for Data Analysis,* New York: John Wiley & Sons, 1977.

APPENDIX A

Optimum Weighting Profile When System 1 Is Operating Over Rayleigh Fading Channels

Consider Weighting System 1 operating with M-fold ideal maximal ratio diversity and NCFSK. The optimum weights that minimize the digital noise power ϵ_a^2 for this system are found by means of the Lagrange multiplier technique,[11] viz:

$$F = F(\phi_1 \cdots \phi_N,\lambda)$$
$$= \sum_{l=1}^{N} A_l \frac{2^{M-1}}{(2+\phi_l\Gamma)^M} + \lambda\left(\sum_{l=1}^{N} \phi_l - N\right), \qquad (78)$$

where λ is a dummy variable. Partially differentiating F with respect to ϕ_l and λ and setting these derivatives to zero we have

$$A_l \cdot \frac{M2^{M-1}\Gamma}{(2+\phi_l\Gamma)^{M+1}} = \lambda. \qquad (79)$$

Using the constraint $\sum_{l=1}^{N} \phi_l = N$, the λ value is given by

$$\frac{1}{N}\sum_{l=1}^{N} (A_l)^{\frac{1}{M+1}} = (2+\Gamma)\cdot\left(\frac{\lambda}{\Gamma M 2^{M-1}}\right)^{\frac{1}{M+1}}. \qquad (80)$$

Using expressions (79) and (80), we arrive at the optimum weights (33).

Note that for the reason of simplicity of analysis we did not impose the physical condition $\phi_l \geqslant 0$. In the solution to the mathematical optimization problem, optimum mathematical weights might be negative for low channel s/n's.

The optimum weights (33) give the minimum digital noise power

$$\epsilon_a^2 = \frac{2^{M-1}}{(2+\Gamma)^M}\cdot\left(\frac{1}{N}\right)^M\cdot\left(\sum_{l=1}^{N} (A_l)^{\frac{1}{M+1}}\right)^{M+1}, \qquad (81)$$

where the optimum weight profile is used at each s/n Γ. For a fixed weight profile, (81) is a lower bound for those Γ's, where the fixed weight profile is not optimized.

The optimum weights for Systems 2 and 3 can also be found by means of the Lagrange multiplier method. In Appendices C and D we have shown the detailed derivations for the Gaussian channel.

APPENDIX B

Optimum Weighting Profile When System 1 Is Operating Over Gaussian Channels

Applying the constraint that the average bit energy is constant, we determine the optimum weights that minimize ϵ_a^2 with the aid of the Lagrange multiplier technique,[11] viz:

$$F = F(\phi_1, \phi_2, \cdots \phi_N, \lambda)$$
$$= \sum_{l=1}^{N} A_l \frac{1}{2} e^{-\phi_l \frac{E}{2N_o}} + \lambda\left[\left(\sum_{l=1}^{N} \phi_l\right) - N\right], \qquad (82)$$

where λ is a dummy variable. Partially differentiating F with respect to the lth weight,

$$\frac{\partial F}{\partial \phi_l} = -A_l \frac{E}{4N_o} e^{-\phi_l \frac{E}{2N_o}} + \lambda; \qquad l = 1, 2, \cdots, N, \tag{83}$$

and upon setting $\partial F/\partial \phi_l$ to zero we have

$$\lambda = A_l \frac{E}{4N_o} e^{-\phi_l \frac{E}{2N_o}}, \tag{84}$$

a constant. Writing eq. (84) as

$$\ln(\lambda) = \ln(A_l) + \ln\left(\frac{E}{4N_o}\right) - \frac{\phi_l E}{2N_o}$$

and after considering all N values of l, we get

$$N \ln(\lambda) = \sum_{l=1}^{N} \ln(A_l) + N \ln\left(\frac{E}{4N_o}\right) - \frac{E}{2N_o} \sum_{l=1}^{N} \phi_l. \tag{85}$$

Applying the constraint of eq. (22) yields

$$\lambda = \frac{E}{4N_o} (A_1 A_2 \cdots A_n)^{1/N} e^{-\frac{E}{2N_o}}. \tag{86}$$

Defining

$$A_o \underset{=}{\Delta} \left[\prod_{l=1}^{N} A_l\right]^{\frac{1}{N}} \tag{87}$$

and equating eqs. (84) and (86), we have the optimum fixed word weighting profile,

$$\phi_l = 1 + \frac{\ln\left(\frac{A_l}{A_o}\right)}{\frac{E}{2N_o}}; \qquad l = 1, 2, \cdots, N. \tag{88}$$

APPENDIX C

Optimum Weighting Profile When System 2 Is Operating Over Gaussian Channels

Applying eq. (26) to eq. (56), and upon using the Lagrange multiplier technique, we have the function

$$F = F(\phi_1(x_0) \cdots \phi_l(x_0) \cdots \phi_N(x_0) \cdots$$
$$\cdots \phi_l(x_i) \cdots \phi_N(x_{2^N-1}), \lambda_0, \lambda_1 \cdots \lambda_l \cdots \lambda_{2^N-1}) \tag{89}$$

$$= \sum_{i=0}^{2^N-1} p_i \sum_{l=1}^{N} a_l(x) \frac{1}{2} e^{-\phi_l(x_i)\frac{E}{2N_o}}$$

$$+ \sum_{i=0}^{2^N-1} \lambda_i \left[\sum_{l=1}^{N} \phi_l(x_i) - N \right]. \tag{90}$$

Upon partially differentiating F with respect to $\phi_l(\mathrm{x}_i)$,

$$\frac{\partial F}{\partial \phi_l(x_i)} = p_i - a_l(x_i) \frac{E}{4N_o} e^{-\phi_l(x_i)\frac{E}{2N_o}} + \lambda_i, \tag{91}$$

and setting $\partial F/\partial \phi_l(x_i)$ to zero we get

$$\frac{\lambda_i}{p_i} = a_l(x_i) \frac{E}{4N_o} e^{-\phi_l(x_l)\frac{E}{2N_o}}, \tag{92}$$

a constant for each fixed i and $l = 1, 2, \cdots, N$. Applying the constraint on $\phi_l(x_i)$, and following the procedure used in the derivation of eq. (88), we have the optimum individual word weighting profile,

$$\phi_l(x_i) = 1 + \frac{\ln\left[\dfrac{a_l(x_i)}{a_o(x_i)}\right]}{\dfrac{E}{2N_o}}$$

$$l = 1, 2, \cdots, N$$

$$i = 0, 1, 2, \cdots 2^N - 1, \tag{93}$$

where

$$a_o(x_i) = (a_1(x_i)a_2(x_i) \cdots a_N(x_i))^{1/N}$$

$$= \left[\prod_{l=1}^{N} a_l(x_i) \right]^{1/N}. \tag{94}$$

APPENDIX D

The Optimum Word Weighting Factor When System 3 Is Operating Over Gaussian Channels

In this case the Langrange multiplier technique yields

$$F = F(W_0, W_1, W_2, \cdots, W_{2^N-1}, \lambda)$$

$$= \sum_{i=0}^{2^N-1} p_i N a_o(x_i) \frac{1}{2} e^{-\frac{W_i E}{2N_o}}$$

$$+ \lambda \left[\sum_{i=0}^{2^N-1} p_i W_i - 1 \right], \tag{95}$$

whence

$$\frac{\partial F}{\partial W_i} = p_i N a_o(x_i) \frac{1}{2} e^{-\frac{W_i E}{2N_o}} \left(-\frac{E}{2N_o}\right) + \lambda p_i.$$

Upon $\partial F/\partial W_i$ equals zero,

$$\lambda = N a_o(x_i) \frac{E}{4E_o} e^{-\frac{W_i E}{2N_o}}; \qquad i = 0, 1, 2, \cdots, 2^N - 1, \qquad (96)$$

a constant. We have assumed that

$$p_i \neq 0; \qquad i = 0, 1, 2, \cdots, 2^N - 1,$$

but if any p_i is zero, the corresponding word weight is set to zero, there being no advantage in reserving energy for an event that never occurs. When some values of p_i are zero the minimization is performed as above, but over those W_i's (fewer than 2^N) that have nonzero p_i.

Taking the natural logarithm of eq. (96) and applying the constraint of eq. (43) yields

$$\ln(\lambda) = \sum_{k=0}^{2^N-1} p_i \ln[N a_o(x_k)] + \ln\left(\frac{E}{4E_o}\right) - \frac{E}{2N_o}. \qquad (97)$$

Thus from eqs. (96) and (97), the optimum word weighting factor becomes

$$W_i = 1 + \frac{1}{\frac{E}{2N_o}} \left\{ \ln[N a_o(x_i)] - \sum_{k=0}^{2^{N-1}} p_k \ln[N a_o(x_k)] \right\}$$

$$i = 0, 1, 2, \cdots, 2^N - 1. \qquad (98)$$

AUTHORS

Raymond Steele, (SM '80), B.S. (Electrical Engineering) from Durham University, Durham, England, in 1959, the Ph.D. degree in 1975, and the Doctor of Science Degree (DSc) in 1983. Prior to his enrollment at Durham University, he was an indentured apprenticed Radio Engineer. After research and development posts at E. K. Cole Ltd., Cossor Radar and Electronics, Ltd., and The Marconi Company, all in Essex, England, he joined the lecturing staff at the Royal Naval College, Greenwich, London, England. Here he lectured in telecommunications to NATO and the External London University degree courses. His next post was as Senior Lecturer in the Electronic and Electrical Engineering Department of Loughborough University, Loughborough, Leics., England, where he directed a research group in digital encoding of speech and television signals. In 1975 his book, *Delta Modulation Systems* (Pentech Press, London), was published. He was a consultant to the Acoustics Research Department at Bell Laboratories in the summers of 1975, 1977, and 1978, and in 1979 he joined the company's Communications Methods Research

Department, Crawford Hill Laboratory, Holmdel, N.J. In 1983 he became Professor of Communications in the Electronics Department at the University of Southampton, England.

Carl-Erik W. Sundberg, M.S.E.E., 1966, and Dr. Techn., 1975, Lund Institute of Technology, University of Lund, Sweden; Bell Laboratories, 1981–1982. Mr. Sundberg is an Associate Professor in the Department of Telecommunication Theory, University of Lund, and a consultant in his field. He is Director of the consulting company SUNCOM, Lund. During 1976 he was with the European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands, as an ESA Research Fellow. He has been a Consulting Scientist at LM Ericsson and SAAB-SCANIA, Sweden, and at AT&T Bell Laboratories. His research interests include source coding, channel coding (especially decoding techniques), digital modulation methods, fault-tolerant systems, digital mobile radio systems, spread spectrum systems, and digital satellite communication systems. He has published a large number of papers in these areas during the last few years. Senior Member, IEEE; member, SER, Sveriges Elektroingenjörers Riksförening.

Wai Choong Wong, B.Sc. (Electronic and Electrical Engineering), Ph.D. (Electronic Engineering), Loughborough University of Technology, Loughborough, England, in 1976 and 1980, respectively. Since April 1980 he has been a Member of Technical Staff in the Communications Methods Research Department of Bell Laboratories at Crawford Hill, Holmdel, New Jersey. His current interest is in speech signal processing, particularly applied to mobile radio environment, digital modulation techniques, and simultaneous transmission of data and speech signals.