

Graceful Degradation of Magnetic Recording Channels

KEES A. SCHOUHAMER IMMINK, SENIOR MEMBER, IEEE

Abstract—Coding techniques are shown to provide an effective means of controlling the effects of bandwidth and gain variations associated with space losses. The combination of a specific channel code and a suitable partial response detection technique is proposed and studied for the purpose of obtaining enhanced robustness. The new technique presented for encoding and decoding digital audio signals offers the advantage of a graceful degradation of the performance when the signal is recorded on a digital recorder with a wide range of bandwidths. This situation may arise, for example, when contact between head and medium is insufficient. The new technique can also be employed for digital video signals or any other information that might carry significance information.

I. INTRODUCTION

THE QUALITY of digitally encoded audio signals is, up to certain system thresholds, almost ideal. If, for some reason, the channel bandwidth is reduced below a certain value, the symbol error rate can quite abruptly become so high that the decoding electronics cannot cope with it. Even when very smart error correction schemes are included, we will find that when the bandwidth is below the channel's threshold value, no useful information can be retrieved by the receiver. This effect is not normally found in analog systems, which usually show a more gradual or graceful degradation.

Bandwidth limitations or loss of high-frequency response of a magnetic recorder may be due to a variety of reasons. One such reason is the increase in spacing loss created by increased spacing that exists accidentally, or purposely, between the active surfaces of the head and the medium. The severity of this spacing loss increases rapidly with decreasing wavelength λ , and can be conveniently calculated by expressing the loss as $54.6d/\lambda$ in decibels [1], where d denotes the head-to-medium spacing. Since the minimum recorded wavelength in current recorders is well below $1 \mu\text{m}$, contact between the head and medium surfaces is of critical significance in magnetic (tape) recording, and is often the principal cause of loss at high information densities.

There are a number of ways in which digital symbols can be represented by physical parameters. They all involve assigning a range of waveforms of a continuously

variable physical function to represent some digital symbol. Most digital systems now in use are binary and synchronous, which means that in each symbol time interval or time slot a condition of, e.g., current or no current, pit or no pit, positive or negative magnetization, etc., is transmitted or stored. (The parlance of the communication engineer is used throughout.) The receiver, under the control of its clock, properly phased with respect to the incoming data, samples the received signal at the middle of each time slot. A single pulse transmitted over a bandwidth-limited system is smeared out in time due to the convolution with the channel's impulse response; a sample at the center of a symbol interval is a weighted sum of amplitudes of pulses in several adjacent intervals. This phenomenon is called *intersymbol interference* (ISI). If the system bandwidth is reduced, the ISI will become more severe. This effect can become so great that the receiver—even in the absence of noise—can no longer distinguish between the symbol value and the ISI and will start to make errors.

In this paper we shall present preliminary results of a new technique [2]–[4] where the ISI has a different effect on the various streams of symbols of different significance. The theory is explained assuming that the transmitted signal originates from a digital audio source, though it is stressed that the new technique can be employed for digital video signals as well, or any other information that might carry significance information. Examples may include control and signaling symbols in a data stream, side information like quantizer slopes, or filter coefficients in more sophisticated voice encoding schemes.

II. BASIC CONCEPT

Since saturation recording is assumed, candidate codewords of length n are chosen with elements from a *binary* alphabet. The codewords are selected in such a way that they divide, according to a predefined frequency domain criterion, the available frequency range of the channel into smaller independent subchannels. The chosen frequency domain criterion will be based on the (Walsh)–Hadamard transform. We now select the codewords, denoted by x , $x_i \in \{-1, 1\}$, $i = 1, \dots, n$, n even, to have an odd parity, that is, the number of $+1$'s is odd. The codeword set is denoted by S . The reason behind this particular choice of code set will become apparent in the following example. It can readily be verified that 2^{n-1} codewords

Manuscript received July 11, 1988; revised January 17, 1989. This paper was presented in part at the 9th Symposium on Information Theory in the Benelux, Mierlo, The Netherlands, May 26–27, 1988.

The author is with Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands.

IEEE Log Number 8927001.

satisfy this condition. We also define the frequency-domain representation y of x , which results from the $n \times n$ Hadamard transform, or

$$y = H_n x$$

where H_n , $H_n(i, j) \in \{-1, 1\}$, is the Sylvester-type $n \times n$ Hadamard matrix. This selection of the frequency-domain transform restricts the codeword length to powers of two ($n \geq 4$). Due to the particular choice of the codeword set S we find (this can easily be verified) for any $x \in S$ that the elements y_i of the vector $y = H_n x$ are nonzero. The basic concept of the new technique will now be illustrated by the following simple example. Suppose, for the sake of simplicity, that the codeword is of length four. The Hadamard matrix is

$$H_4 = \frac{1}{2} \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \end{bmatrix}.$$

The coefficients of the Hadamard matrix are ordered in correspondence with increasingly rapid occurrence of the zero crossings (the usual notion of frequency in the case of the Fourier transform). The $2^{n-1} = 8$ codewords with odd parity are $(+1, -1, -1, -1)$, $(-1, +1, -1, -1)$, $(-1, -1, +1, -1)$, $(-1, -1, -1, +1)$, and their inverse counterparts. Table I shows the eight source words, their corresponding channel representations x , according to a coding rule to be explained shortly, with their associated frequency-domain representations $y = H_4 x$ and the vector z with elements $z_i = 0$ if $y_i < 0$ else $z_i = 1$.

In this example we find that $m = 3$ source symbols can be mapped onto $n = 4$ channel bits, so we conclude that the rate of this code is $3/4$. The new technique is illustrated by assuming that the physical channel has an undesirable response at the low-frequency end. For example, the received signal is superimposed on an unknown (quasi) dc component which may result from baseline wander. Suppose the source word "100" is stored. The corresponding channel representation found from Table I, namely, $(-1, 1, -1, -1)$, is transmitted. It is tacitly assumed that the symbols of the codewords are transmitted serially. The retrieved codeword, denoted by r , generally distorted by ISI and corrupted with noise, is now Hadamard-transformed in the receiver into $n = 4$ frequency components, denoted by \hat{y}_i . Or

$$\hat{y} = H_4 r. \quad (1)$$

The estimated values of z_i , $1 \leq i \leq n$, denoted by \hat{z}_i , are subsequently determined from the polarity of \hat{y}_i , i.e., $\hat{z}_i = 0$ if $\hat{y}_i < 0$ else $\hat{z}_i = 1$. Element \hat{z}_1 is skipped, the other three decoded symbols $\hat{z}_2, \dots, \hat{z}_4$ are the received (estimated) source symbols. It is now easy to understand why this set of codewords has been chosen: detection can be kept very simple by just taking the signs of the frequency transform components \hat{y}_i . The reader may have noticed in

TABLE I
AN EXAMPLE OF A RATE 3/4 DC-INSENSITIVE CODE

| Source | x | y | z |
|--------|-------------|-------------|------|
| 000 | -1 +1 +1 +1 | +1 -1 -1 -1 | 1000 |
| 001 | -1 -1 +1 -1 | -1 -1 -1 +1 | 0001 |
| 010 | -1 -1 -1 +1 | -1 -1 +1 -1 | 0010 |
| 011 | +1 -1 +1 +1 | +1 -1 +1 +1 | 1011 |
| 100 | -1 +1 -1 -1 | -1 +1 -1 -1 | 0100 |
| 101 | +1 +1 +1 -1 | +1 +1 -1 +1 | 1101 |
| 110 | +1 +1 -1 +1 | +1 +1 +1 -1 | 1110 |
| 111 | +1 -1 -1 -1 | -1 +1 +1 +1 | 0111 |

Table I that the source word assignment to a specific codeword has been prepared in such a way that the symbols z_2, \dots, z_4 (the three symbols furthest to the right in the column furthest to the right) are equal to the three source symbols (the column furthest to the left). The variable \hat{z}_1 , which actually contains the unknown dc term of the received codeword, is not used. Rows 2, 3, and 4 of the Hadamard matrix have an equal number of plus and minus terms, so that any unknown superimposed dc term will be canceled here, and will not affect the decoding quality.

Fig. 1 shows a block diagram of a receiver that can be employed when the codeword length $n = 4$. It can be seen that the receiver comprises a bit and word synchronization, and four parallel finite impulse response filters with a response equal to a row of the Hadamard matrix. The polarity of the outputs are taken every four channel bits.

An alternative way of looking at the properties of the new technique is as follows. Let D denote the unit-delay operator corresponding to one modulation or channel symbol interval. Let further the transfer function of the i th subchannel, corresponding to the i th row of the Hadamard matrix, be denoted by $F_i(D)$, so that we simply find

$$\begin{aligned} F_2(D) &= (1 + D - D^2 - D^3)/2 \\ &= (1 - D)(1 + 2D + D^2)/2 \\ F_3(D) &= (1 - D - D^2 + D^3)/2 \\ &= (1 - D)(1 - D^2)/2 \\ F_4(D) &= (1 - D + D^2 - D^3)/2 \\ &= (1 - D)(1 + D^2)/2. \end{aligned} \quad (2)$$

These simple equations show that we have created three parallel partial-response subchannels, and we further notice that the three transfer functions have the factor $(1 - D)$ in common; two subchannels have the factor $(1 - D^2)$ in common.

Partial-response detection techniques with transfer function $(1 - D)$ or $(1 - D^2)$, later denoted by the shorthand notation PR(1, -1) and PR(1, 0, -1), respectively, have been described or proposed in experimental digital video recording systems [5]-[8]. Essentially, the channel waveform extends over a small number of symbol intervals in a well-defined manner, which thus causes an

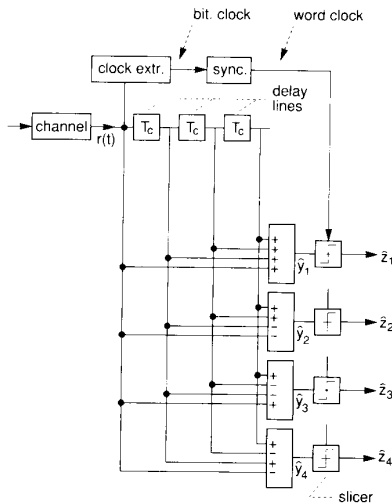


Fig. 1. Block diagram of receiver.

increased number, usually three, of amplitude levels at the discriminator point. It is well known that the performance of these simple detection techniques may suffer from amplitude variations of the retrieved signal since the signal assumes, for the simple partial-response systems studied, three levels at the discrimination point. In order to prevent error propagation, a simple change-of-state precoder step is usually performed at the transmitter's end [5]. Difficulties in the implementation of a "reliable" automatic gain control circuit, in particular during and immediately after dropouts when it is difficult to maintain the correct signal levels, have prevented its wide acceptance in recording practice. The quintessence of the new concept is that by virtue of the particular choice of the codebook (Table I), the signals at the discriminator point in the new format assume *two* levels, which has its bearing on the detection quality when the system is perturbed by amplitude and/or bandwidth variations. A change-of-state precoder step is not required in the new format. The new format is inherently robust against amplitude variations of the received signal since only the polarity of the various signals is taken. An obvious disadvantage of the new technique is the rate loss incurred.

At this point, it is illustrative to show the transfer functions of the various subchannels. The transfer function of the i th, $i = 2, 3, 4$, subchannel $F_i(fT_c)$, where T_c denotes the channel bit length, can easily be found from (2) by substituting $D = e^{-i2\pi fT_c}$. Let, for example, $i = 2$, then

$$\begin{aligned} F_2^2(fT_c) &= \frac{1}{4} |1 + e^{-i2\pi fT_c} - e^{-i4\pi fT_c} - e^{-i6\pi fT_c}|^2 \\ &= 16 \sin^2 \pi fT_c \cos^4 \pi fT_c. \end{aligned}$$

In a similar fashion, we find

$$F_3^2(fT_c) = 16 \cos^2 \pi fT_c \sin^4 \pi fT_c$$

and

$$F_4^2(fT_c) = 4 \sin^2 \pi fT_c \cos^2 2\pi fT_c.$$

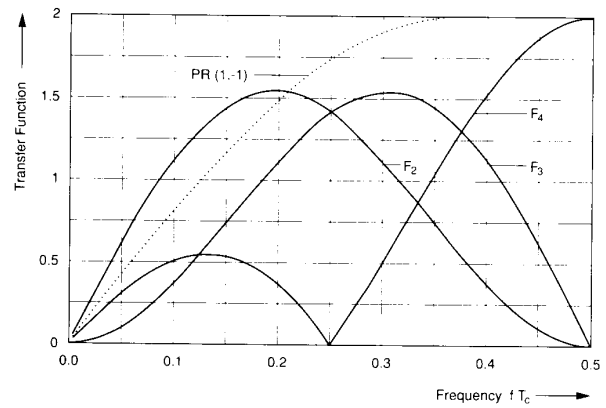


Fig. 2. Transfer function of the subchannels versus frequency. The dotted line is the transfer function of the partial-response system with response $(1 - D)$, the same user bit rate is assumed as in the coded case.

Fig. 2 shows the three transfer functions along with the transfer function of a conventional partial-response system $(1 - D)$ with the same user information rate as the coded system. The plot of the transfer functions reflects what we intuitively expect, namely, the available frequency band is divided into three subbands; the higher the number of the subchannel, the higher the maximum of the transfer function is situated in the frequency band. In recording channels with severe attenuation of the high frequencies, we may expect, after equalization which boosts the higher frequencies, additive noise to be less dominant in the lowest subchannel. Therefore, symbol errors caused by additive noise are more likely to occur in the "higher" subchannel. Naturally, in a system involving storage of ordered data such as digital audio, or any other information that might carry significance information, more significant information is placed in the most reliable subchannel that is in the "lowest" subchannel. Quantitative results that support these arguments will be provided in the next section.

In a typical embodiment of the receiver, the magnetic recorder is equalized in such a way that the recorder followed by the equalizer corresponds with a partial response $PR(1, -1)$ channel. The equalizer output is forwarded to three (note that the channel corresponding to \hat{z}_1 is skipped) parallel finite impulse response filters with response $(1 + 2D + D^2)$, $(1 - D^2)$, and $(1 + D^2)$, respectively, in a similar way as depicted in Fig. 1. Though the concept can be generalized to an 8×8 Hadamard matrix and a rate $7/8$ code, for reasons of clerical convenience, results will be provided for the simple rate $3/4$ code only. Attempts to find other transforms that constitute the given constraints and unique decodability have, up till now, failed.

This introductory section has sought to outline in qualitative terms the main idea of the frequency-domain detection technique; in the following section we will take a closer look at the more quantitative effects of intersymbol interference and additive noise.

III. CHANNEL MODEL

In this section a comparison is made, using a simple, but not unrealistic, model of the magnetic recorder, between the more commonly used partial-response detection technique and the new coding technique. It is assumed that the binary user information with a bit rate of $1/T$ is translated into a coded channel sequence having a channel bit rate of $1/T_c$, $T_c \leq T$. The quotient $R = T_c/T$ is, as usual, called the *rate* of the code. The recorded sequence \mathbf{a} consists of binary digits $a_i \in \{-1, 1\}$ that are generated each T_c seconds. The recorded channel sequence $\mathbf{a} = (a_1, \dots, a_M)$, taken to be of arbitrary length M , is a member of a predefined set S_c of codewords. Assuming a linear readout mechanism, the retrieved signal $r(t)$ is of the form

$$r(t) = \sum_{i=1}^M a_i g(t - iT_c) + n_w(t) \quad (3)$$

where $g(t)$ is the channel waveform and $n_w(t)$ is additive white Gaussian noise with two-sided spectral density $N_0/2$. In other words, it is assumed that noise originates solely from the read electronics. The retrieved signal is low-pass-filtered and subsequently sampled at $t = kT_c$.

The analysis of the magnetic recorder channel is based on the Lorentzian channel model [9]. According to this model, the step response of the readout process is of the form

$$h(t) = A \frac{1}{1 + (2vt/pw_{50})^2} \quad (4)$$

where v is the medium-to-head speed, pw_{50} determines the dispersivity of the recording channel, and A is a constant of proportionality. The magnetic domains are assumed to be recorded as perfect full- T_c pulses $p_{T_c}(t)$ of unity amplitude, or

$$\begin{aligned} p_{T_c}(t) &= 1, & |t| < \frac{T_c}{2} \\ &= 0, & \text{otherwise.} \end{aligned} \quad (5)$$

The channel waveform $g(t)$ of the combined recording in conjunction with the readout process is given by

$$g(t) = h(t + T_c/2) - h(t - T_c/2). \quad (6)$$

The corresponding channel transfer function is

$$G(fT_c) = A\pi S \sin(\pi fT_c) e^{-\pi|f|T_c S/R} \quad (7)$$

where the normalized information density S is defined as

$$S = \frac{pw_{50}}{vT}. \quad (8)$$

This definition is given on the understanding that the user bit rate $1/T$ is constant in both the coded and uncoded situation.

The performance of the new system is evaluated assuming that a zero-forcing equalizer precedes the Hadamard transform. The technique of the (zero-forcing) equalizer

and its performance are well established so that we will not dwell on the details [10]. The performance of the new technique when an equalizer is employed in cascade with the Hadamard transform can be obtained in the following way.

Let σ_i^2 denote the normalized noise variance in the i th subchannel at the decision (discrimination) point. By integrating the power spectral density of the noise at the decision point, we obtain an expression for the noise variance σ_i^2

$$\sigma_i^2 = \frac{E_b}{N_0} \frac{1}{R} \int_0^{1/2} \frac{F_i^2(fT_c)}{G^2(fT_c)} dfT_c \quad (9)$$

where $F_i(fT_c)$ is the transfer function of the i th, $i = 2, 3, 4$, subchannel, and

$$E_b = A^2 \pi^2 T$$

denotes the energy-per-user bit. In the simple case under consideration, we have $T_c = 3/4T$ and $R = 3/4$. The noise variance of the conventionally, uncoded, that is, $R = 1$, partial response PR(1, -1) and PR(1, 0, -1) detected signal is found by substituting

$$F_i^2(fT_c) = 4 \sin^2 \pi fT$$

and

$$F_i^2(fT_c) = 4 \sin^2 2\pi fT$$

respectively. Hence

$$\sigma_{\text{PR}(1,-1)}^2 = \frac{E_b}{N_0} \int_0^{1/2} \frac{4 \sin^2 \pi fT}{G^2(fT)} dfT \quad (10)$$

and

$$\sigma_{\text{PR}(1,0,-1)}^2 = \frac{E_b}{N_0} \int_0^{1/2} \frac{4 \sin^2 2\pi fT}{G^2(fT)} dfT. \quad (11)$$

Essentially, the former results coincide with those derived by Osawa *et al.* [7], [8]. The probability of receiving a symbol in error in the i th subchannel is denoted by $\Pr(e_i)$, $i = 2 \dots 4$.

For moderate and high signal-to-noise ratios, the error probability $\Pr(e_i)$ is well approximated by the following expression:

$$\Pr(e_i) \approx Q\left(\sqrt{\frac{E_b}{N_0} \frac{1}{\sigma_i}}\right) \quad (12)$$

where $Q(x)$ is the tail of the Gaussian distribution, defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} dy$$

and the noise variances σ_i^2 are easily found from (9).

IV. ROBUSTNESS AGAINST SPACING LOSS

Up till now we have concentrated on the resistance of the new and conventional schemes to additive white noise. The conventional analysis that considers errors resulting

from additive random noise is not sufficient, and it is therefore important to take into account other effects that may reduce the system's performance. One such effect is the increase in spacing loss created by decreased head-to-medium contact [7]. The spacing loss approximates the classic Wallace equation [1]

$$L_s \approx e^{-(2\pi d/\lambda)} \quad (13)$$

where L_s is the spacing loss, d is the spacing (relative to the nominal value) that exists accidentally, or purposely, between the active surfaces of the head and the medium, and λ is the recorded wavelength. Since the minimum recorded wavelength in current recorders is well below 1 μm , contact between the head and medium surfaces is of critical significance in magnetic recording, and is often the main cause of loss at high information densities. The reduction of the playback amplitude introduced over tens and hundreds of symbol intervals is considered to be the result of an increase of head-to-medium spacing [11]. Combination of (7) and (13) yields the following transfer function $\tilde{G}(fT_c, \delta)$ of the recorder:

$$\begin{aligned} \tilde{G}(fT_c, \delta) &= L_s G(fT_c) \\ &\approx A\pi S \sin(\pi fT_c) e^{-\pi|f|T_c R^{-1}(S+2\delta)} \end{aligned}$$

where the normalized medium-head spacing δ is defined by

$$\delta = \frac{d}{vT}$$

In general, the minimum recorded wavelength is reduced in coded streams by a factor R , assuming that both the coded and uncoded systems achieve the same user density, so that, obviously the coded system becomes more vulnerable to spacing loss. Exceptions are codes based on run-length-limited sequences [12], sometimes termed λ -increasing codes, which have an increased resistance to spacing loss.

The overall sampled transfer function of the i th subchannel during channel bandwidth mismatch is

$$\tilde{F}_i(D) = F_i(D) (\cdots, \beta_{-1}D^{-1}, 1 - \beta_0, \beta_1 D, \cdots)$$

where the residual intersymbol interference terms β_k are found from

$$\beta_k = 2 \int_0^{1/2} \cos 2\pi k f T_c (e^{-2\pi f T_c \delta / R} - 1) df T_c \quad (14)$$

When the medium-head spacing is small, i.e., $\delta \ll 1$, we find

$$\beta_0 \approx \frac{\pi}{2} R^{-1} \delta$$

and

$$\beta_1 \approx \frac{2}{\pi} R^{-1} \delta.$$

The channel waveform error, as a result of the bandwidth mismatch, is simply reflected in a reduction of the amplitude $1 - \beta_0$ of the retrieved signal and residual intersymbol interference terms β_k , $k \neq 0$.

When the residual intersymbol interference terms β_k , $|k| > 1$, are ignored, we find for the error probability of the various subchannels during mismatch [3]

$$\begin{aligned} \Pr(e_2) &\approx K_2 Q\{(1 - \beta_0)(1 - \beta_1)/\sigma_2\} \\ \Pr(e_3) &\approx K_3 Q\{(1 - \beta_0)(1 - 2\beta_1)/\sigma_3\} \\ \Pr(e_4) &\approx K_4 Q\{(1 - \beta_0)(1 - 3\beta_1)/\sigma_4\} \quad (15) \end{aligned}$$

where $K_i \approx 1/4$ are constants, not written out in full here for the sake of brevity.

Equations (15) are very basic for the technique developed here. The effects of noise and intersymbol interference can easily be understood. We observe that the lower a particular subchannel is situated in the frequency band of the channel, the less it is affected by residual intersymbol interference, or, in other words, the more rugged it is against impaired reception due to receiver mismatch.

By way of comparison, we note that the sampled transfer function of conventional PR(1, -1) detection is of the form

$$\begin{aligned} \tilde{F}_{\text{PR}(1,-1)}(D) \\ = (1 - D) (\cdots, \beta_{-1}D^{-1}, 1 - \beta_0, \beta_1 D, \cdots). \end{aligned}$$

The worst case eye opening can be found after a routine computation

$$\begin{aligned} \text{eye}_{\text{PR}(1,-1)} = \min \left\{ \tilde{F}_{\text{PR}(1,-1)}(0) - \tilde{F}_{\text{PR}(1,-1)}(1) - 1 \right. \\ \left. - \sum_{i \neq 0,1} |\tilde{F}_{\text{PR}(1,-1)}(i)|, \tilde{F}_{\text{PR}(1,-1)}(0) \right. \\ \left. + \tilde{F}_{\text{PR}(1,-1)}(1) - \sum_{i \neq 0,1} |\tilde{F}_{\text{PR}(1,-1)}(i)| \right\} \end{aligned}$$

so that, after some evaluation, we find the error probability during mismatch

$$\Pr(e_{\text{PR}(1,-1)}) \approx Q\left\{ \frac{1 - 2(\beta_0 + 2\beta_1)}{\sigma_{\text{PR}(1,-1)}} \right\} \quad (16)$$

where it has been tacitly assumed that the threshold levels are held fixed at their nominal value. No attempt has been made to assess the performance when an automatic threshold control is incorporated. This choice is indeed a matter of debate, and the reader is warned at this point that conclusions on the detection quality during mismatch, with or without the assistance of such a device, may be considerably different [11].

The preceding relationships are best illustrated by an actual numerical example. We consider here the rate $R = 3/4$ code, and assume a relative information density $S = 1.5$. Then, for the coded case we find $\beta_0 \approx 2.1\delta$, $\beta_1 \approx$

0.85 δ and for the uncoded case $\beta_0 \approx 1.57\delta$, $\beta_1 \approx 0.63\delta$. A substitution, using (15) and (16), shows that

$$\begin{aligned}\Pr(e_2) &\approx K_2 Q\{(1 - 3.0\delta)/\sigma_2\} \\ \Pr(e_3) &\approx K_3 Q\{(1 - 3.75\delta)/\sigma_3\} \\ \Pr(e_4) &\approx K_4 Q\{(1 - 4.75\delta)/\sigma_4\}\end{aligned}$$

and

$$\Pr(e_{\text{PR}(1,-1)}) \approx Q\{1 - 5.9\delta\}/\sigma_{\text{PR}(1,-1)}.$$

Thus we may conclude that in this particular case none of the three subchannels created is more seriously affected by intersymbol interference than the equivalent uncoded partial response channel. Another result which is of interest is the fact that when intersymbol interference becomes so severe that the conventional PR(1, -1) channel closes its eye, i.e., when the argument of the Q function vanishes, two out of the three subchannels are still in good working condition, which actually shows that the new detection technique offers the great advantages of graceful degradation and robustness. Strictly speaking this cannot be derived from the previous relations that are valid for small δ . A computer search which takes into account virtually all residual intersymbol interference was used to confirm this statement.

In a similar fashion, it is possible to derive the worst case eye opening of partial response PR(1, 0, -1) detection.

V. AUDIO QUALITY

It is not possible to compare the performance of the newly developed format, which basically consists of three independent subchannels, and the conventional system. The performance measure depends too heavily on the particular application. When, for example, the transmission of control data and signaling symbols of a hierarchic system is based on the new format, then the performance measure becomes extremely complex.

When dealing with a performance measure for digital sound it is notoriously difficult to specify quantitatively the degree of annoyance experienced by a human observer. This holds in particular for digital sound signals that are source-encoded. For straight PCM-encoded signals, Dostis [13] arrived at the following simple mathematical relation that takes into account both the effects of quantization and incorrect decoding of the received data.

$$\text{SNR} = \frac{1}{2^{-2N} + 4 \sum_{i=1}^N \Pr(E_i) 2^{-2i}} \quad (17)$$

where SNR is defined as the signal-to-noise ratio of the reconstructed audio signal, $\Pr(E_i)$ is the probability that the i th audio symbol is erroneously received. It is assumed that the sound is linearly quantized with 2^N levels and that a natural representation is employed. When the

channel is in good condition, i.e., $\Pr(E_i)$ is small, we find the well-known relation between the SNR and the number of quantization levels

$$\text{SNR} = 2^{2N} \approx (6N) \text{ dB}.$$

An experiment on the effectiveness of the new technique is to compare its performance with that of the conventional, uncoded, digital data transmission system. For purposes of illustration we have computed the SNR of an $N = 15$ -bit linear quantization system based on a system with an $n = 4$ codeword length as a function of the information density. The choice of the value of N is not arbitrary. We choose the codeword assignment in such a way that five codewords accommodate a 15-bit audio sample. The $3 \times 5 = 15$ bits that constitute an audio sample are transformed into $4 \times 5 = 20$ channel bits. It is evident that the most significant audio bits require the best protection against channel errors; therefore, we assume a data format where the five most significant bits of the audio sample are placed in the lowest, that is, the most reliable, subchannel, etc. We now find that the probability $\Pr(E_i)$ of erroneously receiving an audio symbol is given by $\Pr(E_i) = \Pr(e_2)$, $1 \leq i \leq 5$, $\Pr(E_i) = \Pr(e_3)$, $6 \leq i \leq 10$, and $\Pr(E_i) = \Pr(e_4)$, $11 \leq i \leq 15$. Recall that in this format the first symbol \hat{z}_1 is skipped and the symbols \hat{z}_2 , \hat{z}_3 , and \hat{z}_4 correspond to the 3-bit inputs from the audio source. In the uncoded case, all audio bits are treated equally and thus have an equal probability of being erroneously received. Using the equations of the previous section, it is now straightforward to compute the audio performance. Fig. 3 shows the audio quality, expressed in SNR of both the new and conventional partial response PR(1, -1), and PR(1, 0, -1) systems, as a function of the information density S , where it is assumed that $10 \log E_b/N_0 = 30$ dB. It should be appreciated that both the uncoded and coded systems are compared at the same user information rate; in other words, the computations are performed assuming $T_c = 3/4T$, which is, of course, the "fairest" way to do. We conclude that the new system performs slightly better than the conventional PR(1, -1) system. It further becomes clear that PR(1, 0, -1) is the best choice. The benefits of the new system become clear when disturbances other than additive noise are present.

It is now assumed that the channel is perfectly equalized when there is no spacing loss, that is, the equalizer is designed for the transfer function $\tilde{G}(fT_c, \delta = 0)$, but the actual transfer function is $\tilde{G}(fT_c, \delta)$. The relative information density is $S = 1.5$, and $10 \log E_b/N_0 = 30$ dB. The performance of the new and conventional schemes is shown in Fig. 4, which relates audio quality to head-to-medium spacing. The range of the head-to-medium spacing depicted is so large that the small signal analysis as previously discussed is not valid, and a computer search has to be used to evaluate the worst case eye opening.

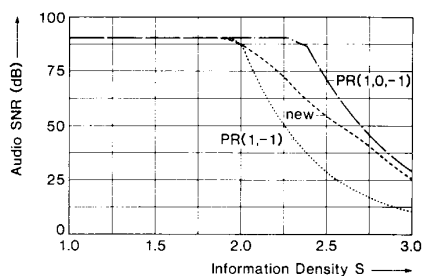


Fig. 3. Audio SNR versus information density S (--- new system, ··· conventional PR (1, -1) system, - · - conventional PR (1, 0, -1) system), $10 \log E_b/N_0 = 30$ dB.

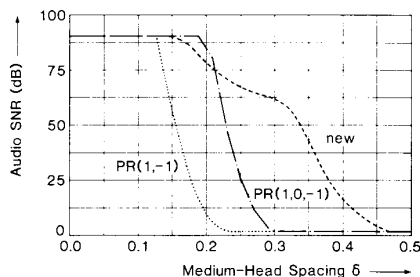


Fig. 4. Audio SNR versus the medium-head spacing (--- new system, ··· conventional PR (1, -1) system, - · - conventional PR (1, 0, -1) system). Relative information density $S = 1.5$, $10 \log E_b/N_0 = 30$ dB.

Looking at Fig. 4, it can be noticed that, in the conventional PR (1, -1) and PR (1, 0, -1) systems, the audio performance rapidly deteriorates with increasing spacing loss. (The PR (1, 0, -1) system is again superior to the PR (1, -1) system.) The new system shows a more graceful degradation with increasing medium-head spacing. Results of computations for other values of the information density S and channel signal-to-noise ratio E_b/N_0 , not shown for reasons of space, show that the behavior as depicted in Fig. 4 is quite typical.

VI. CONCLUSIONS

We have provided some preliminary results demonstrating the potential, and quite substantial, performance advantages in the use of coding to combat the effects of bandwidth and gain variations associated with space losses. The combination of a specific channel code and a suitable partial response detection technique has been proposed and studied for the purpose of obtaining enhanced robustness. The new technique presented offers the virtue of a graceful degradation. Intersymbol interference has a different effect on the error probability of the least and most significant symbols. The class of systems described may well be of particular value in applications involving storage of ordered data such as digital audio, or any other information that might carry significance information, when the impulse response is strongly time-variant owing, for example, to spacing loss. Examples may include

control and signaling bits in a data stream, side information like quantizer slopes, or filter coefficients in more sophisticated voice or audio encoding schemes. While our modeling assumptions are somewhat idealistic, nevertheless, the results reported here are encouraging and hopefully will provide some impetus for further investigations in this area.

ACKNOWLEDGMENT

The author wishes to thank Prof. S. Tazaki, Ehime University, for helpful discussions.

REFERENCES

- [1] R. L. Wallace, "The reproduction of magnetically recorded signals," *Bell Syst. Tech. J.*, vol. 30, pp. 1145-1173, Oct. 1951.
- [2] K. A. S. Immink, "Graceful degradation of digital audio transmission systems," *J. Audio Eng. Soc.*, vol. 36, no. 1-2, pp. 27-33, Jan./Feb. 1988.
- [3] —, "Transmission of digital sound over channels with unforeseen bandwidth limitations," *J. Inst. Electronic Radio Eng.*, vol. 58, no. 2, pp. 49-56, Mar./Apr. 1988.
- [4] —, "Graceful degradation of digital sound reproduced from magnetic recording channels," in *Proc. 9th Information Theory Symp. in the Benelux* (Mierlo, The Netherlands), pp. 33-39, May 1988.
- [5] H. Kobayashi and D. T. Tang, "Applications of partial response channel coding to magnetic recording systems," *IBM J. Res. Develop.*, vol. 14, pp. 368-375, July 1970.
- [6] R. W. Wood and D. A. Petersen, "Viterbi detection of class IV partial response on a magnetic recording channel," *IEEE Trans. Commun.*, vol. COM-34, pp. 454-461, May 1986.
- [7] H. Osawa, Y. Okamoto, and S. Tazaki, "Performance comparison of partial-response systems in digital VTR," *Electron. Commun. Japan*, pt. 1, vol. 68, pp. 75-83, July 1985.
- [8] H. Osawa, S. Tazaki, and S. Ando, "Performance analysis of partial-response systems for non-return-to-zero recording," *IEEE Trans. Magn.*, vol. MAG-22, pp. 253-258, July 1986.
- [9] C. D. Mee and E. D. Daniel, *Magnetic Recording*. New York, NY: McGraw-Hill, 1987.
- [10] J. G. Proakis, *Digital Communications*. New York, NY: McGraw-Hill, 1983.
- [11] S. Nakagawa, K. Yokoyama, and H. Katayama, "A study on detection methods of NRZ recording," *IEEE Trans. Magn.*, vol. MAG-16, pp. 104-110, Jan. 1980.
- [12] P. H. Siegel, "Recording codes for digital magnetic storage," *IEEE Trans. Magn.*, vol. MAG-21, no. 5, pp. 1344-1349, Sept. 1985.
- [13] I. Dostis, "The effect of digital errors on PCM transmission of compressed speech," *Bell Syst. Tech. J.*, vol. 44, pp. 2227-2243, 1965.

Kees A. Schouhamer Immink (M'81-SM'86) was born in Rotterdam, The Netherlands, on December 18, 1946. He received the B.S. degree from the Rotterdam Polytechnic in 1967 and the M.S. and Ph.D. degrees from the Eindhoven University of Technology in 1974 and 1985, respectively, all in electrical engineering.

He joined the Philips' Research Laboratories, Eindhoven, in 1968, where his work involved the signal processing side of optical recording systems. He was, among others, responsible for the design and development of channel coding techniques for the compact disc, compact disc video, and experimental erasable optical audio discs. Since 1986, he has been a senior scientist of the magnetic recording group at the Philips' Research. He has written numerous papers in the field of coding techniques for optical and magnetic recording, and is co-author of a book entitled *Principles of Optical Disc Systems*. He was awarded a Fellowship by the Audio Engineering Society in 1985 for "his work in the area of optical laser disc and for detailed study on channel codes for the compact disc." He holds more than thirty patents, mainly in the area of optical recording.

Dr. Immink is a Fellow of the IEE.