The Roles of Multi-Level Signalling and Multi-Valued Logic in Digital Communications

J.J. Bergin*

K.C. Smith**

* Telecommunications Engineer, Ontario Hydro, Toronto, Canada.
**Department of Electrical Engineering, University of Toronto, Toronto, Canada.

ABSTRACT

A review is presented of multi-level signalling in conventional communications systems. Possibilities are introduced for the application of multi-valued logic (MVL) in digital modulation schemes for data transmission.

A general overview of digital baseband modulation techniques is given first. Modulation schemes that involve binary to multi-level conversions and multi-level coding are then discussed, and their advantages and disadvantages outlined.

This paper does not attempt to present a complete MVL communications system, but rather provides an overview from which possible applications of MVL circuits to the communications area can be appreciated. Its goal is to direct attention to the relevance of recent developments in MVL circuits to communications systems applications.

1.0 OVERVIEW

That MVL has many potential applications in the broad field of communications has already been noted. For example, at the chip level, MVL can alleviate the pin limitation problem that characterizes VLSI technology. [24,10] Presently, the pin limitation constraint leads to the necessity to time-share binary signals on data highways and I/O pins, which in turn involves speed restraints, testing difficulties and other problems. Multi-valued signalling allows each input pin to accept (and each output pin to deliver) more information. For an equivalent information transfer the total number of pins required in an integrated circuit chip utilizing MVL signalling can be far fewer than needed in a chip utilizing simple binary techniques.

As well, as will be shown, in the digital modulation area there is aleady a general tendency toward multi-level* signalling in order to increase the information rate over a fixed bandwidth channel. Already, for this application, MVL circuits are potentially useful for the level conversion from binary, if required, and for the elaborate coding schemes usually incorporated.

2.0 COMMUNICATIONS BASICS

Most data communication systems can be represented by the block diagram of Figure 1. Communication in the voice band whether over twisted wire pairs, long haul or short haul over the Bell switching network, or within private networks, can be modeled in this manner. However, if data is transmitted in a higher frequency range, for example, microwave, then Figure 1 would be representative of only one data channel, which must be frequency multiplexed with a number of others before being transmitted.

*Note that multi-valued, as used in the MVL literature, has its origin in mathematics, whereas multi-level, as used in communications, describes a partially equivalent idea originating in a practical domain in which voltage levels are assigned discrete values (e.g. -3V, -2V, -1V, 0, +1V, +2V, +3V in a 7-level system).

Generally speaking, the encoder will convert the input binary bit stream, usually NRZ (non return to zero) encoded, into a multi-level signal and insert some form of error detecting/correcting code. Error detecting and correcting codes will be discussed later.

An understanding of the next stage, that is, filtering, will give an appreciation of the major problem of data communications systems, namely intersymbol interference (ISI).

2.1 FILTERING

Signal shaping and filtering are used in data communications systems to confine the signal to a specific frequency band and to control ISI.

Application of Nyquist theory [5, 11, 14] to minimum bandwidth transmission systems (refer to Figure 2a and 2b) shows that it should be possible to transmit f_s independent symbols in a channel constrained (as in a low-pass filter) to have a bandwidth of only $f_1=f_s/2$.

Not only is the infinite slope cutoff of the Nyquist filter difficult to achieve because of the infinite number of sections which are needed to produce the required sharp cutoff, but also its $\sin x/x$ type of impulse response (Figure 2b) exhibits an undesirable slow rate of decay. One problem with this slow decay is that the resulting system has practically no tolerance to error in sampling times or other system perturbations.

The goal then is to approximate the Nyquist channel using a filter with "raised-cosine roll-off" characteristics as shown in Figure 2c where the measure of rolloff is given as a percent.

With present-day technology, an approximation utilizing 30% filters is feasible, requiring only a 30% excess Nyquist bandwidth. In a normalized channel (f_1=1 Hz) having 0% rolloff factor, it is theoretically possible to transmit at a rate of 2 symbols per second, corresponding to a rate of 2 bits/s/Hz for binary transmission systems. If the rolloff factor equals 30%, then the transmission rate is approximately 1.54 bits/s/Hz.

When a baseband signal with the shaping indicated in Figure 2c and 2d is used to amplitude modulate a carrier, both sidebands of the resulting double sideband signal have the same rolloff characteristics (Refer to Figure 3).

Control of spectral shaping can be accomplished either in the baseband before modulation or in the passband after modulation, or by a combination of both.

Once the overall transmission characteristics are chosen to minimize ISI there remains the problem of the best division of the characteristics between the transmitting and receiving filters. In this process it is common to try to maximize the signal-to-noise ratio (S/N) at the demodulator, assuming white Gaussian noise and a fixed limit on transmitted power.

0195-623X/85/0000/0115$01.00 © 1985 IEEE

2.2 CHANNEL CAPACITY

Transmission efficiency was introduced in section 2.1. It will now be discussed in more detail.

Bandwidth and noise power place a restriction on the rate of information that can be transmitted by a channel [15]. It can be shown that in a channel contaminated by white Gaussian noise, one can transmit information at a rate of, at most, 'C' bits per-second, where 'C' is the channel capacity given by:

$$C = B\log_2(1 + S/N) \quad (1)$$

Here, B is the channel bandwidth in Hz, S is the signal power, and N is the noise power. Equation 1 is known as the Hartley-Shannon law and is considered to be the central theorem of information theory. It is directly valid for white Gaussian noise but can be modified for other types of noise.

Basically the equation is derived as follows: If the root mean square value of the received signal is $(S+N)^{1/2}$ volts and the root mean square value of the noise voltage is $(N)^{1/2}$ volts, it follows that the number of distinct levels that can be distinguished without error will be given by:

$$M = \frac{(S+N)^{1/2}}{(N)^{1/2}} = (1 + S/N)^{1/2}$$

The maximum amount of information carried by each pulse, or symbol, having $(1+S/N)^{1/2}$ distinct levels is given by:

$$I = \log_2(1 + S/N)^{1/2} \text{ bits} = 1/2\log_2(1 + S/N) \quad (2)$$

An ideal Nyquist filtered system with bandwidth B Hz can transmit a maximum of $2B$ pulses, or symbols, per second. Since each pulse can carry the maximum information specified by equation 2, it follows that a system of bandwidth B can transmit information at a rate of:

$$C = B\log_2(1 + S/N) \text{ bits per second} \quad (1)$$

The capacity 'C' of a channel (or system) is thus limited by the bandwidth and the noise which characterizes it.

This channel capacity 'C' represents the maximum amount of information that can be transmitted by a channel per second. To achieve this rate of transmission the information has to be processed properly or coded in an efficient manner.

It is evident that the bandwidth and signal power can be exchanged for one another. For example, consider a transmission of pulses at a rate of $2f_m$ samples per second in a channel whose bandwidth is f_m Hz, where each of the samples can assume any of 16 states. Therefore $(S+N)^{1/2}/(N)^{1/2} = 16$. In order to transmit $2f_m$ pulses per second, we need an ideal Nyquist channel bandwidth of f_m Hz. The channel capacity for a 16 level signal in this channel is:

$$C = f_m\log_2((S+N)/N)$$
$$= f_m\log_2(16)^2$$
$$= 8f_m \text{ bits per second}$$

Now consider the signal to noise ratio required at the receiver to distinguish pulses that assume only four distinct states. It is $(S+N)^{1/2}/(N)^{1/2} = 4$. This mode of transmission would allow the required signal power to be reduced for a given value of M and a fixed value of N. However, we now have to transmit twice as many pulses per second, namely $4f_m$, to retain the same information rate. Therefore, the channel bandwidth required is $2f_m$ Hz, for which the channel capacity is:

$$C = 2f_m\log_2((S+N)/N)$$
$$= 2f_m\log_2(4)^2$$
$$= 8f_m \text{ bits per second, as before}$$

It should be mentioned that the process of exchange of bandwidth and signal power is not automatic. The signal information would have to be transformed (coded) or modified in some way to occupy the desired bandwidth. In practice, this is accomplished by various types of modulation. Some modulation forms prove superior to others in utilizing the channel capacity.

It should be emphasized that M-ary* transmission has a higher spectral efficiency than a corresponding communication system using binary pulses. If, in a binary case, there are $2f_m$ pulses per second sent over an ideal Nyquist channel, with bandwidth f_m, then the bandwidth efficiency would be, $2f_m/f_m$ information bits/s/Hz, which is equal to 2 b/s/Hz, since each pulse represents one information bit. However, if each symbol, or pulse, transmitted contains 'n' information bits, where $n = \log_2 L$, and 'L' equals the number of levels the signal may take on, then the efficiency would be $2f_m n/f_m$ information bits/s/Hz=2n b/s/Hz. Figure 4 gives an indication of spectral efficiency for various modulation schemes.

2.3 EQUALIZATION

An equalization stage is required in radio and voice frequency systems to mitigate intersymbol interference (ISI) caused by the fading channel phenomenon, as well as imperfect filtering, amplification, modulation, etc. [3,8,25] The equalizinion function can be accomplished at either the passband or in the baseband after demodulation.

Equalization becomes more complex, for a given bandwidth, with increased data rate. At passband, equalization efforts have concentrated mainly on amplitude impairments. In very efficient systems (having high transmission rates in small bandwidths (M-ary systems)) more is done to correct for delay impairments. In the last several years there has been considerable interest in using baseband transversal equalizers to correct for both amplitude and delay deficiencies.

2.4 MODULATION

The type of modulation used in data communications will depend on the combined effects on performance of interference, fading, and delay distortion. In addition, the modulation schemes, and encoding arrangements as well, must be compared with respect to spectral efficiency (the data transfer rate for a given bandwidth), bit error rate (BER) signal to noise ratio**, and system complexity. [17]

There are three basic modulation techniques commonly used: amplitude (AM), frequency (FM) and phase (PM). Each of these basic techniques has a large number of variations. In addition, a hybrid, amplitude-phase shift keying (APK), is receiving some attention. [22] Refer to Figure 4 for a performance comparison between some of the basic modulation schemes.

The simplest digital AM technique is double sideband (DSB) amplitude modulated by a binary signal. The waveform is represented as:

$$x(t) = A/2(1 + m(t))\cos w_c t$$

where $m(t)$ is the modulating signal and w_c is the carrier frequency. For the case of 100% modulation by a non-return-to-zero (NRZ) binary data waveform (for which $m(t) = \pm 1$), we have an on-off keying modulation (OOK).

Since the carrier conveys no information, efficiency can be improved by the use of double sideband suppressed carrier (DSB-SC) AM.

$$x(t) = Am(t)\cos w_c t$$

*In M-ary systems M represents the number of separately distinguishable signals than can be detected at the receiver. The number of different amplitude levels the signal may take on is some function of M.

**S/N is sometimes substituted with E_b/N_0, which is the average signal energy/bit to noise power spectral density ratio as measured at input to the receiver.

Quadrature amplitude modulation (QAM) is an AM alternative. This technique involves summing two DSB-SC signals 90 degrees apart in phase:

$$x(t) = A(m_I(t)\cos w_c t + m_Q(t)\sin w_c t)$$

When $m_I(t)$ and $m_Q(t)$ are three-level signals (+1,-1, or 0), called duo-binary, coded in such a way as to minimize ISI caused by filtering. The result is quadrature partial response (QPR), modulation, which has been proposed for use in the Canadian 8 GHz frequency band, and will be mentioned later. These QAM techniques require coherent detection. Any phase-tracking errors that occur, result in interference between the $m_I(t)$ and $m_Q(t)$ signals, thereby degrading performance.

When $m_I(t)$ and $m_Q(t)$ take on values ± 1, QAM is identical to quaternary phase-shift keying (QPSK) discussed shortly in the paragraph on PM techniques.

M-ary or M-level QAM is now coming into use. It is based on the same principle as QAM but $m_I(t)$ and $m_Q(t)$ can now take on M/4 levels. The advantage of this will be discussed later.

FM techniques involve frequency modulating or frequency shift keying (FSK) the carrier. As with other modulation schemes, FSK can be detected either coherently or non-coherently. Continuous phase FSK (CP-FSK) is a coherent scheme that has received some attention. [18,19] Abrupt phase changes at the bit transition instants characteristic of other FSK forms are avoided. Memory, imposed on the waveform by the requirement for continuous phase transitions improves performance, reducing error probability for a given S/N ratio. This improvement results from the use of several symbols to make a decision rather than the more common approach of making independent symbol-by-symbol decisions.

Another FM technique is Minimum Shift Keying, MSK [1]. This technique achieves the same general performance as coherent FSK with the superior spectral properties of CP-FSK. It also has the additional advantage of a relatively simple self-synchronizing implementation, which coherent CP-FSK does not have.

In digital PM schemes the carrier phase is shifted by the modulating data stream. For binary phase shift keying (BPSK) the shift is either 0 or 180 degrees. Detection requires a precise phase reference. Quaternary PSK (QPSK) schemes have also been developed. Coherent QPSK involves encoding two bits at a time into one of four possible carrier phases spaced 90 degrees apart. As in the QAM case, M-ary PSK schemes have been developed where $\log_2 M$ bits at a time can be encoded into M fixed carrier phases.

At the receiver, baseband information is recovered by a detection process. Coherent detection requires a sinusoidal reference signal perfectly matched in both frequency and phase to the received carrier. This phase reference may be obtained either from a transmitted pilot tone or from the modulated signal itself. Noncoherent detection, based on waveform characteristics, is independent of phase, but uses instead, energy or frequency, and does not require a phase reference.

Usually, detection is followed by a decision process, incorporated in a decoder, that converts the recovered baseband modulation signal into a sequence of binary digits. This process requires bit synchronization information, which is generally extracted from the received waveform. With most modulation schemes, decisions can be made on a bit-by-bit basis with no loss in performance. However with some schemes, an advantage can be gained by examining the signal over several bit intervals prior to making each bit decision. The portion of the received waveform examined by the decision device is called the *observation interval*.

2.5 CODING

One final point should be made about communication systems in general. This concerns the role of error detecting and error correcting codes.

In a practical system there are occasional errors, and it is the purpose of codes to detect and perhaps correct them. These codes cannot correct every conceivable pattern of errors but rather must be designed to correct only the most likely patterns. Much of coding theory has been based on the assumption that each symbol is affected independently by the noise, so that the measure of an error pattern is only the number of errors it contains. For example, codes have been developed that correct any pattern of 't' or fewer errors in a block of 'n' symbols. While this may be an appropriate model for some channels, it is not for others: For example, on telephone lines and on magnetic storage systems, errors occur predominantly in bursts. Consequently, codes for correcting bursts of errors are required, and some good codes have been developed [12].

Block codes [5,12] of which there are many types, have been used extensively. The encoder for a block code breaks the continuous sequence of information digits into n-symbol sections or blocks. It then operates on these blocks independently according to the particular code to be employed. With each block of 'n' data symbols there is a k-tuple of checking symbols, where 'n' is greater than 'k'. The code word is transmitted corrupted by noise, and decoded independently of all other code words. At the receiver, a decision is made on the basis of the information in the received k-tuple, concerning the validity of the code word transmitted.

One simplified coding technique is to use the k-tuple digits to send the number of 1's or 0's in the n-sequence code word. This could help to detect when an error has occurred in the code word, and a request to retransmit would be sent by the receiver. An obvious disadvantage of error detecting/correcting codes is the fact that the information rate decreases due to the additional error detecting symbols in the transmission and the requirement to retransmit data.

If it can be determined by the error detecting circuitry that a certain symbol is in error, and the system is a binary one, then the error can easily be corrected. (If it is not '0' then it must be a '1'). However, error correcting becomes more challenging when dealing with a multi-level system, where the symbols in the data stream can take on a number of values at any given time.

Convolutional codes [5], also known as sequential or recurrent codes, differ from block codes in that the check bits are continually interleaved in the coded bit stream rather than being grouped into words. The encoding/decoding procedure therefore is a continuous process, eliminating the buffering or storage hardware required with block codes.

It is been shown that elaborate coding formats can improve system performance [9]. (That is, lower the bit error rate). A systematic search of binary convolutional codes, has shown [9] that for a bit error probability of less than 2×10^{-4}, the saving in the average (power) signal to noise ratio over that required for an uncoded system with the same channel symbol rate was more than 2dB with PSK and APK, and 1.5dB with AM modulation schemes.

2.6 SUMMARY

The digital communications field is so large that it is difficult to cover all aspects in one paper. This overview has attempted to give only an appeciation of what is involved with filtering, coding and digital modulation systems.

The type of modulation used will basically depend on the tradeoffs between performance, bandwidth usage, and equipment complexity (which usually means decoder complexity) that one is willing to make. For example, convolutional coding will improve system performance at the expense of an increase in required bandwidth. M-ary QAM and PSK schemes, by expanding the signal sets, have a good spectral efficiency but also degrade the performance somewhat.

Superior communication systems have been achieved by combining the convolutional coding techniques with an expanded signal set. This has the net effect of no bandwidth expansion but improved performance.

In addition, a further improvement can be made by introducing a form of memory to the system. For example in CP-FSK, continuous phase modulation which utilizes memory, improves the spectral characteristics of the signal transmitted since the phase transitions can be made to occur more smoothly. Also, correlative coding, or partial response coding which will be discussed in detail later, is another technique that incorporates memory, forcing the decoder to examine the received signals over a longer span of time. This has the effect of decreasing the required bandwidth for a given information rate.

In the following sections M-level signalling, and partial response signalling and coding, that have been mentioned previously, will be treated in more detail. It is in these techniques and also in Level Division Multiplexing that potential applications of multi-valued logic circuits will be shown to exist.

3.0 MULTI-LEVEL SIGNALLING

3.1 M-ARY QAM AND M-ARY PSK

As illustrated in Figure 4, modulation schemes having the highest spectral efficiency are M-ary PSK and M-ary QAM. Both schemes have been studied extensively for this reason. Multi-amplitude or multi-phase signalling requires no greater bandwidth than binary to increase the information rate over the binary case, whereas, multi-tone FSK does.

As a specific example of the spectral efficiency of an M-ary system, consider a 16-ary QAM signal. As Figure 5 indicates, after the 2-to-4 level conversion, the signal rate is given as $f_S = (f_b/2)(1/log_2 4)$, where f_b is the incoming binary data rate. Since $f_b = 10Mb/s$ and $M = 16$, then $f_S = 2.5Mb/s$. Therefore, if the low pass filter at the next stage were ideal, with 0% rolloff, then the minimum bandwidth required would be 1.25 MHz, according to the Nyquist theory. After the modulation stage the bandwidth requirement for the Bandpass filter (BPF) would be twice the baseband spectrum, or 2.5 MHz, as implied by Figure 3. Therefore, the spectral efficiency is 10 Mb/s/2.5 MHz = 4 b/s/Hz. Practical, high speed, 40 Mb/s 16-ary QAM systems have achieved a bandwidth efficiency of approximately 3.7 b/s/Hz.

The disadvantage coming with increased spectral efficiency is, as Figure 4 indicates, a necessary increase in signal power to maintain the same bit error rate (BER). In addition, there is need for more complex circuitry to encode and decode the multilevel signals. As well, the high number of bits per symbol will require increased performance from the system components, such as filters and adaptive equalizers [3].

Despite these disadvantages, M-ary PSK and QAM are coming increasingly into use. As circuits are simplified and integrated, costs are decreasing and, as more elaborate encoding techniques are being developed, the BER can be decreased even further for a given spectral efficiency. Further, it seems that recent developments in multi-valued logic circuits can be a contributing factor to the increased use of multi-level signalling in the data communications area, as the reader will see shortly.

To reach an appreciation of the difference between PSK and QAM, the signal space diagrams, or 'signal constellations', for 16-aray PSK and 16-ary QAM are shown in Figure 6. Each point in signal space represents the amplitude and phase of the received signal. It can be seen that in PSK schemes, only the phase is varied and the received signal does not change in amplitude, whereas in 16-ary QAM, it does.

Figure 7 illustrates the basic process involved in the transmitting stage of an 8-PSK digital radio channel.* It shows that even though there is no amplitude variation in the transmitted signal, there is a point at which an input binary data stream (rate = f_b) must be level-converted to a quaternary signal. (Refer to Table 1)

This may be one area where recent advances in multi-valued logic toward VLSI implementation could help reduce the chip area required by level conversion. W. Current [7] has introduced a voltage-mode encoder/decoder circuit for binary to quaternary logic conversion. His encoder uses an NMOS transistor voltage divider string and a single power supply to set up four reference voltages which are easily converted to the required voltage levels indicated in Table 1. The simulated delays for the NMOS MVL circuit, approximately 200 ns, would be sufficient for voice frequency modem applications, with a bit rate less than 20 kb/s, and perhaps for some digital radio applications having bit rates less than 10 Mb/s.

Figure 8 illustrates the receiver setup for either M-ary QAM or M-ary PSK, where the obvious requirement is shown for an M-ary to binary converter to restore the multi-level signal to its original binary form.

Table 1

Signal Level at 'a'	Logic Level at 'A'	'C'
+1.307	1	1
+ .541	1	0
- .541	0	0
-1.307	0	1

3.2 PARTIAL RESPONSE TECHNIQUES

Another application of multi-level signalling is in 'partial response',** or 'correlative', coding techniques for spectral shaping. There are two basic methods for shaping the signal power spectrum. One method, already discussed, is by filtering the waveform (in the frequency domain), the other is by coding the signals (in the time domain).

Partial response coding is the process of converting the original two-level signal into (typically) a three-level signal and introducing a controlled level of ISI. Figure 9 illustrates the effect of three-level partial response signalling in the time domain.

Partial response coding/filtering can be modelled as in Figure 10a. This block diagram can be redrawn as a single 'conversion filter', shown in Figure 10b which would replace the filtering stage of Figure 1.

The conversion filter can be used to process binary or multi-level signals, as shown in Figure 10c.

The encoder is included to prevent a phenomenon called error propagation, which occurs when a received error causes additional errors in succeeding symbols due to the inherent symbol correlation. At 'A_t', the data appears in a non-binary form and has 2^n levels. The symbols are uncorrelated and each represents two or more binary digits. The encoder converts the Q-level

* Reference 11, chapter 6

**Reference 11, chapter 7

input, consisting of independent digits, into a Q-level source with a memory extending over a fixed number of digits, (typically one or two). Next, level conversion is accomplished by using a level conversion filter. Such a filter causes an overlap of pulses and introduces a controlled amount of ISI. As a consequence, the number of levels at 'C_t' is 2Q-1. Correlation properties are inherent in such a waveform. In addition, it is possible to associate each level with the non-binary input at 'A_t'. At the receiving end, the reconverter reconstructs the original data input at 'A_t' by identifying each digit independently, without resorting to the past history of the correlative waveform. In this way received errors do not cause additional errors in the succeeding symbols.

One specific encoding and conversion process for the Q-level case is:

$$A_t = B_t + D^2 B_t \quad (mod\ Q) \tag{1}$$

$$C_t = B_t - D^2 B_t \quad \text{(algebraic)} \tag{2}$$

$$C_t modQ = A_t \tag{3}$$

where D^2 corresponds to a two symbol delay. At the receiver, A_t is determined by simply performing the operation $C_t modQ$.

The error detection process is a little more involved. One advantage of correlation systems compared to zero memory systems is their error-detection capability. Error detection in zero memory systems requires a certain amount of redundancy in the symbols sent. Correlation systems have finite memory and this can be utilized to monitor and detect errors without introducing redundant digits at the transmitter.

Figure 11 is the general diagram for error detection for binary or non-binary correlative systems. The incoming waveform, with (2Q-1) levels is C_t: the 'binary slicers' determine (by thresholding) which of the 2Q-1 levels was sent: the 'Logic and Sampling' stage converts the C_t levels into the original A_t symbols. ($\hat{A}_t = A_t$ plus errors).

Error detection is done by first determining whether the extreme levels (top and bottom) correspond to the present and past digits originating from the encoder and digital memory. A comparison is made at the sampling instants of the extreme level digits of C_t. If there is a disagreement an error is indicated and the memory is reset to the correct state. A comparison is done when the extreme levels are present because only the extreme levels are formed in a unique way in correlative systems. Intermediate levels may be formed in more than one way and are therefore not suitable for the detection of errors. Figure 12 provides an example of the error detecting process for $Q=4$.

A particular MVL logic gate, the Cycling Gate [6,16], performs the function:

$$x^y = (x + y) \quad modR$$

with R, the radix. It seems that there are definite applications for MVL cycling gates, where R is greater than 2, in the design of partial response schemes, since equation 3 is a modulo addition (where Q is the radix).

In addition, the quaternary NMOS high threshold detector, shown in Figure 13 [21], has possible applications in the partial response error detecting process. If $Q=4$ in the partial response receiver of Figure 11, then the circuit of Figure 13 could easily act as the top slicer, which in this particular case would detect the '3' level.

The partial response technique, as a modulation scheme, has been used extensively since its development in the early 1960's.

For example, the long-haul microwave network in Canada employs a baseband 5800 kHz wide to carry some 1200 voice frequency channels. These are usually grouped together in the band above 564 kHz by means of frequency division multiplexing (FDM). The space below 564 kHz, which is normally not utilized, has been considered for data transmission.

In fact, it was shown to be possible to insert a 1.544 Mb/s data signal below 564 kHz. This was accomplished by utilizing partial response coding. The 1.544 Mb/s data rate was reduced in half by taking the strictly digital binary signal and converting it to a quaternary signal, which reduced the data rate to 772 kHz. Partial response coding/filtering, complete with error detecting coding, reduced the rate to approximately 500 kHz. The complete technique is called 'Data Under Voice' (DUV) [4].

3.3 CODING

With the increased usage of multi-level signalling there is a corresponding demand for multi-level coding techniques. Ungerboeck [23] addresses this in a recent paper. He describes a coding technique which improves error performance of synchronous data links without sacrificing data rate or requiring more bandwidth. This is achieved by channel coding with expanded sets of multi-level/multi-phase signals. The codes can be interpreted as binary convolutional codes with a mapping of coded bits into multi-level channel signals by 'set partitioning'.

A simplified multi-level technique that is in use now is the assigning of binary digits in a Gray code format to the multi-level symbols. For example, if a symbol is incorrectly interpreted by the receiver as being one threshold level above the correct one then the converted binary word will only have one bit in error. That is, if each symbol represents a multiple-bit binary word, then consecutive symbol levels should be assigned binary words that differ by only one bit.

One advantage in transmitting code words composed of symbols that can take on more than two values is that the 'distance' between each code word can be made large, and thus there are inherently more code words to utilize for a given code word length. One 'distance' measure is called the Hamming distance [12], which is defined as the number of bit positions in which the code words differ. Therefore, for a multi-level data signal there will be more code words that satisfy some minimum distance requirement, or any other coding distance requirement imposed on the code words by the encoder.

There is a general mathematical relationship between distinguishability of code words and the power needed to transmit them reliably. The relation first formulated by Shannon, was discussed previously, and although many signalling systems have been constructed, schemes that perform as well as Shannon's theory promises have still not been found.

One way to design a signalling system that comes close to meeting the standards of Shannon's theorem is to represent each signal (code word) as a point in n-dimensional space. There is a technique called 'sphere packing' [5,20], which addresses this problem, but is quite involved and will not be discussed further.

One possible multi-level coding technique would be to restrict the level transitions of the transmitted data. Consider a purely digital PAM (pulse amplitude modulation) system where the transmitted signal is a sequence of amplitude modulated pulses with the number of possible pulse levels, L, greater than two. Normally, each pulse may take on any of the L possible values. However, if for each pulse level the next available level is restricted to lie within a constrained 'distance', then the pulse sequence that results would have smoother transitions than an uncoded pulse sequence. This would have the effect of lowering the bandwidth of the digital signal transmitted. The encoder would have to be designed such that it assigned only code words or sequences that followed this maximum transition rule. Individual code words in the set could differ greatly, but the encoder would have to ensure that the transitions between separate code words or sequences also obeyed the maximum transition rule.

It seems that as the need for processing of multi-level signals increases in data communications there will be a logical increase in the demand for circuits that can perform multi-valued processing efficiently.

3.4 LEVEL DIVISION MULTIPLEXING

Another area where multi-valued logic circuits may have application is in level division multiplexing (LDM) [13]. LDM is a technique that supplements a typical PAM signal with additional information, called a service channel, by utilizing the extreme levels of the PAM signal. This approach does not produce any worsening of the transmission quality of the information in the main channel.

Additional information is usually added to a channel by means of time multiplexing (TDM) or frequency multiplexing (FDM). However, both these methods are expensive and require an increase in the channel bandwidth.

For LDM the service channel is added according to Figure 14a. The bits of the supplementary stream are stored in a buffer and sent to the level generator or the transmitter one at a time, when an extreme level in the main information signal is transmitted. In this case the level generator must be capable of generating also the supplementary levels. The occurence of an extreme level of the main information signal is detected by a simple logic circuit connected to the coder output. At the receiver the decision circuit will have supplementary decision thresholds. The decoding of the main information is done in the normal manner but an additional decoder recovers the supplementary bits which must be buffered to provide an output stream with a constant rate. Refer to Figure 14b and 14c.

The random manner in which the service channel bits arrive at the receiver and the obvious possibility of error propagation are solved by an elaborate arrangement in which the supplementary information exists in a frame of its own, independent of the frame of the main information signal, with each frame containing a proper alignment word. The beginnings of the frames and of their intervals are independent of the occurrence of the extreme levels of the main information signal. They are fixed at the transmitter by a cyclic counter fed by pulses at the same rate as the symbols of the main information signal (11.456 MHz). At the receiver, in the steady state, an identical counter recovers the frame timing.

4.0 CONCLUSION

It has been shown that there are potential applications of multi-valued logic circuitry in digital communications.

With an increase in the use of M-ary PSK and M-ary QAM techniques there is a corresponding increase in the demand for simple multi-level/binary encoder/decoders. That is, the requirement is for MVL converters that are easily integratable since there is a trend in the data communications field to integrate complete transmitter arrangements on one microprocessor chip. While there are many possible relatively low frequency applications, for microwave communications there is also a requirement for MVL converters that are capable at high data rates (in excess of 100 Mb/s).

A general trend in data communications has been to attempt to improve system performance by incorporating a form of memory to the signal, by developing more elaborate coding techniques, and by expanding the signal sets to multi-level. It is believed that current developments in MVL circuitry may find application in these areas. Such applications include decoding and error detecting multi-level signals that have a certain amount of correlation between symbols in the observation intervals, and directly implementing the MV arithmetic required for MV coding schemes. As well, it is hoped, the availability of MVL circuits may encourage the development of more complicated multi-level codes.

5.0 Acknowledgements

The authors gratefully acknowledge the support of the Canada's Natural Sciences and Engineering Research Council through NSERC grant A1753.

6.0 BIBLIOGRAPHY

[1] Amoroso, F., Pulse and Spectrum Manipulation in the Minimum (Frequency) Shift Keying (MSK) Format, IEEE transactions on Communications, VOL. COM-24, No.3, March 1976, pp.381-384.

[2] Anderson, C., Barber, S., Modulation: The Key to Effective Bandwidth Utilization, Telesis, Dec. 1977, pp.172-179.

[3] Aprille, T., A Survey of Networks for Digital Radio, IEEE International Symposium on Circuits and Systems, May 1983, pp.839.

[4] Baker, D., Baart, J., DG1: Equipment for Parallel Data Under Voice Systems, Telesis, Summer 1975, pp.46-52.

[5] Carlson, A., *Communication Systems*, McGraw Hill Inc., USA 1975, Chapter 10.

[6] Carmona, J., et al., Realization of Three-valued CMOS Cycling Gates, Electronic Letters, 27th April 1978, Vol. 14, No.9, pp.288-290.

[7] Current, W., et al., VLSI Chip Interfaces with Quaternary Logic, IEEE International Symposium on Circuits and Systems, May 1983, pp.520.

[9] Davey, J., Modems, Proceedings of the IEEE, Vol. 60, No.11, Nov. 1972, pp.1284-92.

[9] Digeon, A. On Improving Bit Error Probability of QPSK and 4-level Amplitude Modulation Systems by Convolutional Coding, IEEE Transactions on Communications, VOL COM 25, No.10, Oct. 1977, pp.1238-39.

[10] Etiemble, D., MV Integrated Circuits for Microcomputers Systems, IEEE 1982 COMPCON Fall, pp.135-144.

[11] Feher, K., *Digital Communications*, Prentice Hall Inc., New Jersey, 1981, Chapter 3.

[12] Gallager, R., *Information Theory and Reliable Communications*, New York, Wiley, 1968, Chapter 1.

[13] Giusto, P., Level Division Multiplexing of Service Channels in Multilevel Digital Transmissions, IEEE Transactions on Communications, VOL COM 27, No.12, Dec. 1979, pp. 1946-52.

[14] Huang, Feher, and Gendron, Techniques to Generate ISI and Jitter-free Bandlimited Nyquist Signals and a Method to Analyze Jitter Effects, IEEE Transactions on Communications, VOL COM 27, No.11, Nov.1979, pp.1700-11.

[15] Lathi, B., *Communication Systems*, Wiley, New York, 1968, Chapter 8.

[16] Mouftah, H., Smith, K., Three-valued CMOS Cycling Gates, Electronic Letters, 19th Jan., 1978, Vol.14, No.2, pp.36-37.

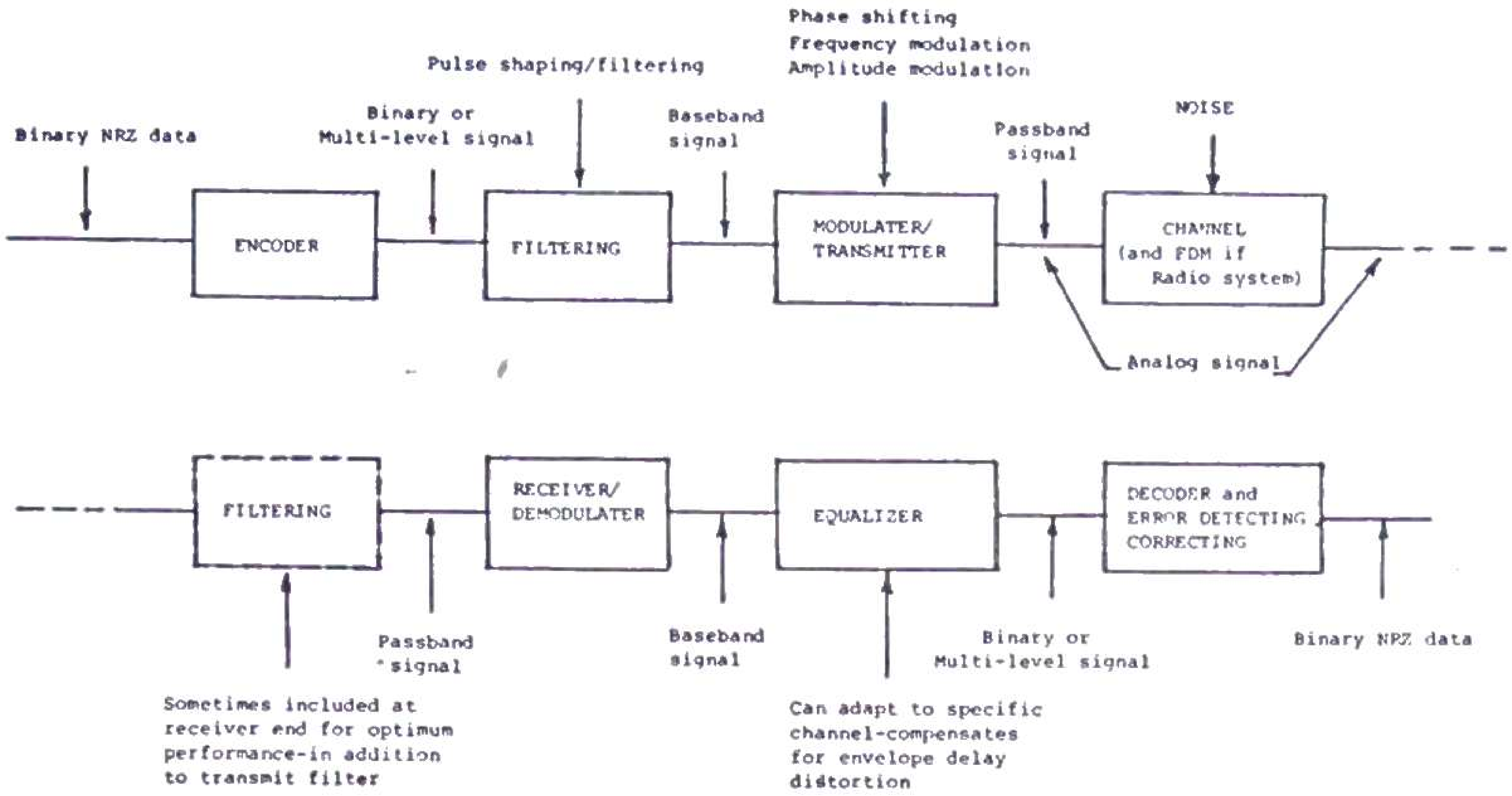


Figure 1

Block diagram showing the stages characteristic of most data communication systems

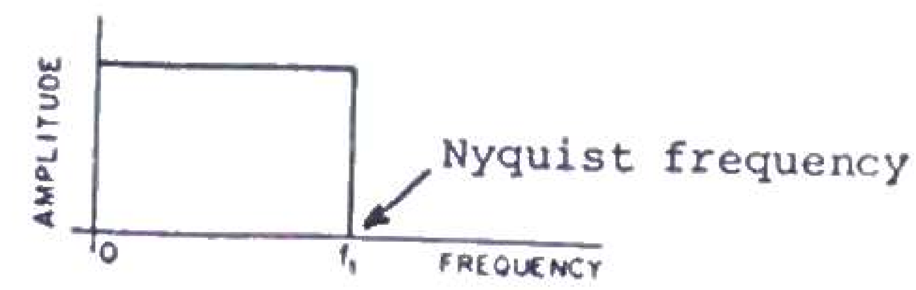


Figure 2a
Ideal Nyquist baseband signal with rectangular spectrum

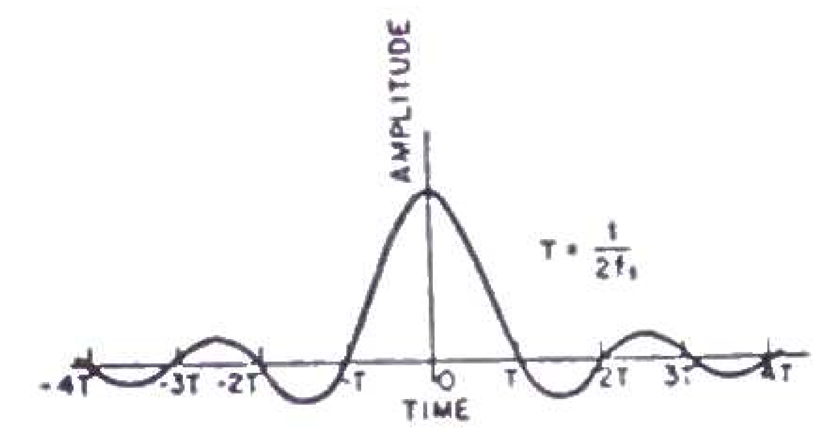


Figure 2b
Time response corresponding to the frequency response of Figure 2a

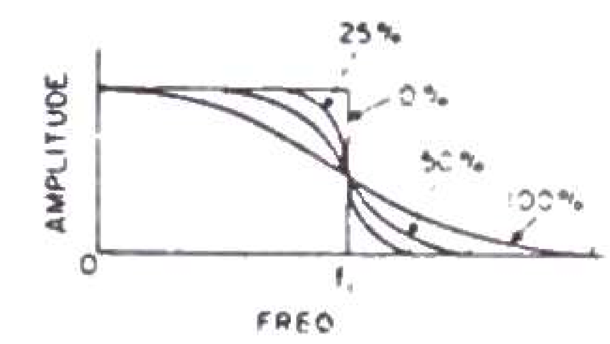


Figure 2c

Possible modified baseband spectra

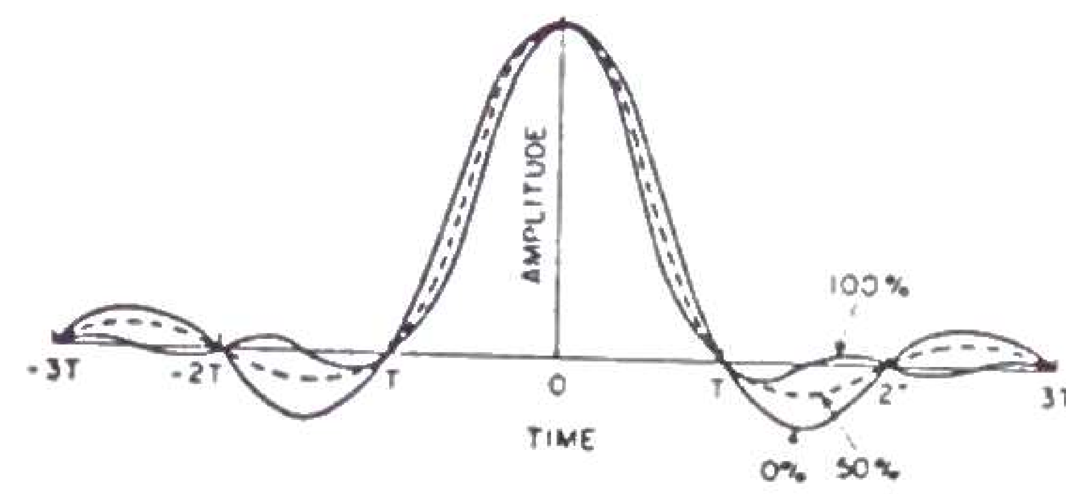


Figure 2d
Time response corresponding to Figure 2c

[17] Oetting, J.,
A Comparison of Modulation Techniques for Digital Radio, IEEE Transactions on Communications, VOL COM 27, No.12, Dec. 1979, pp.1752-62.

[18] Osborne, W., and Luntz, M.,
Coherent and Noncoherent Detection of CPFSK, IEEE Transactions on Communications, VOL COM 22, No.8, Aug. 1974, pp.1023-36.

[19] Schonhoff, T.,
Symbol Error Probabilities for M-ary CPFSK: Coherent and Noncoherent Detection, IEEE Transactions on Communications, June 1976, pp.644-46.

[20] Sloane, N.,
The Packing of Spheres, Scientific American, Jan. 1984, pp.116-125.

[21] Smith, K.,
The Prospects for Multi-Valued Logic: A Technology and Applications View, IEEE Transactions on Computers, Vol. C30, No.9, Sept. 1981, pp.619-634.

[22] Thomas, Weidner and Durrani,
Digital Amplitude-Phase Keying with M-ary Alphabets, IEEE Transactions on Communications, VOL COM 22, No.2, Feb. 1974.

[23] Ungerboedk, G.,
Channel Coding with Multi-Level/Phase Signals, IEEE Transactions on Information Theory, Vol. IT-28, No.1, Jan. 1982, pp.55-67.

[24] Vranesic, Z.,
Applications and Scope of Multi-Valued LSI Technology, IEEE 1981 COMPCON, Spring, 1981.

[25] Wong, W., Greenstein, L.,
Multipath Fading Models and Adaptive Equalizers in Microwave Digital Radio, IEEE Transactions on Communications, VOL COM 32, No.8, Aug. 1984, pp.928-34.

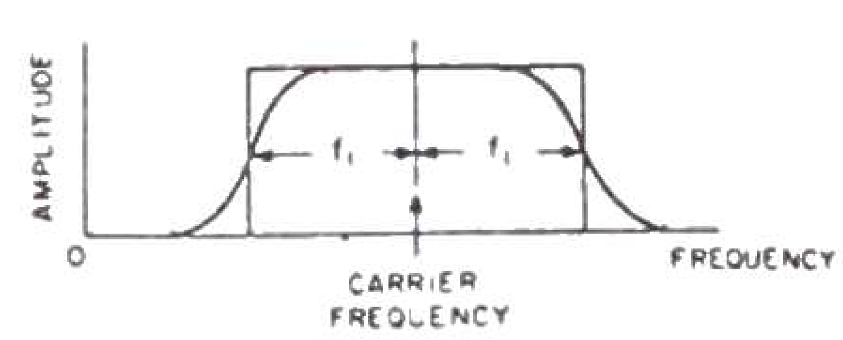


Figure 3

Baseband spectrum shape after amplitude modulation

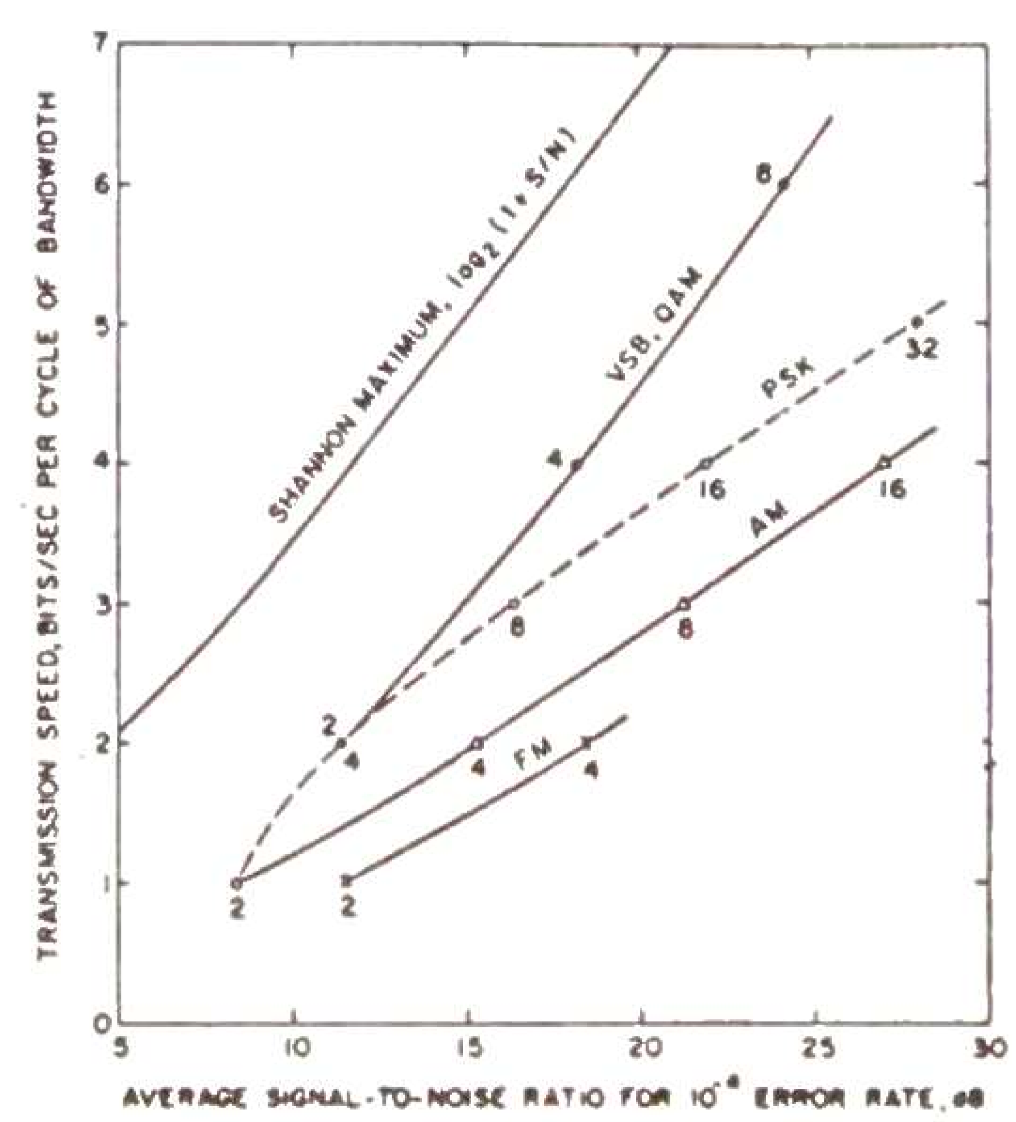


Figure 4
(Reference 8, pg. 1288, Fig. 8)

Spectral effeciency vs S/N for various modulation schemes. Numbers on curves indicate the value of M for M-ary schemes.

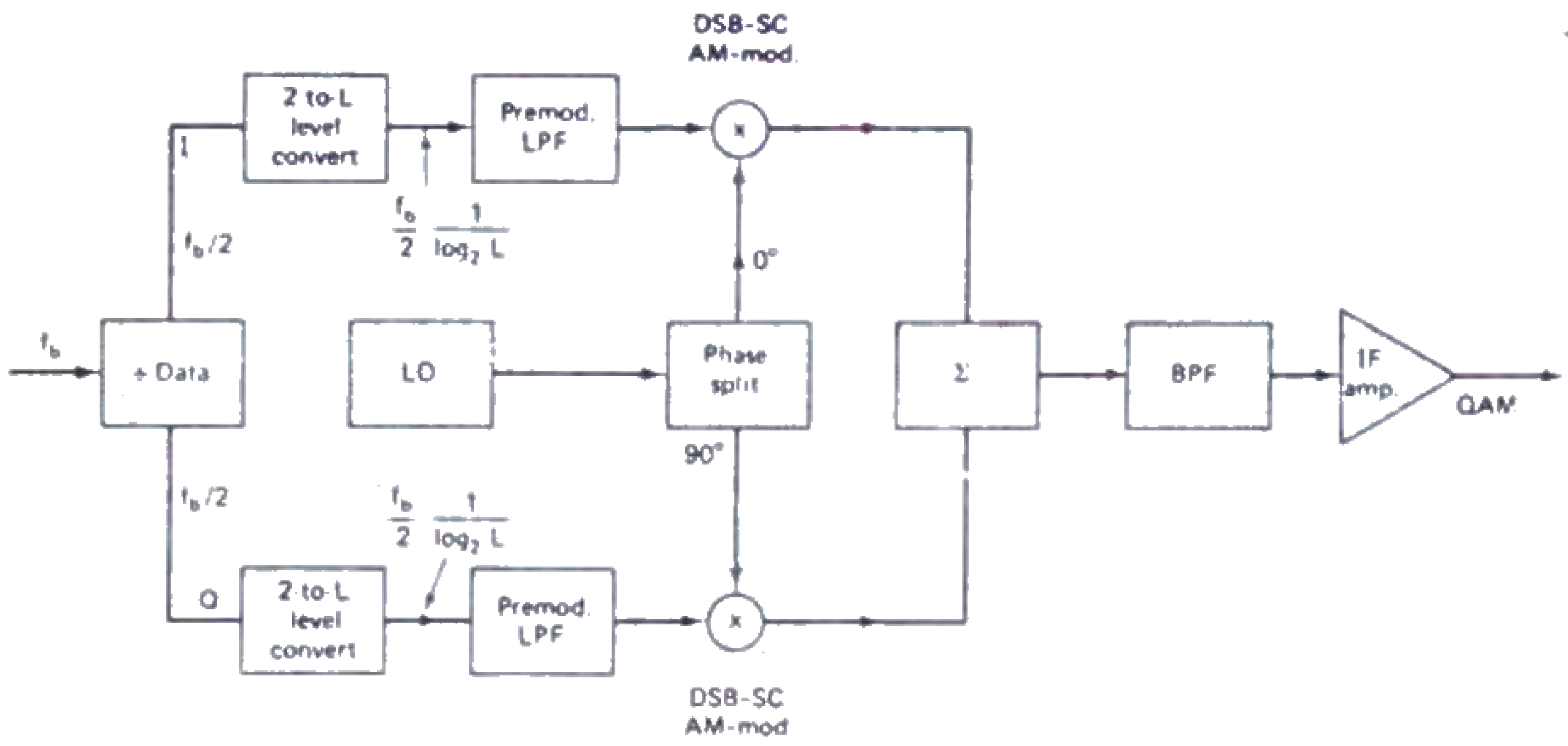


Figure 5: (Reference11, pg. 124, Fig. 6.9)

M-ary QAM modulator block diagram

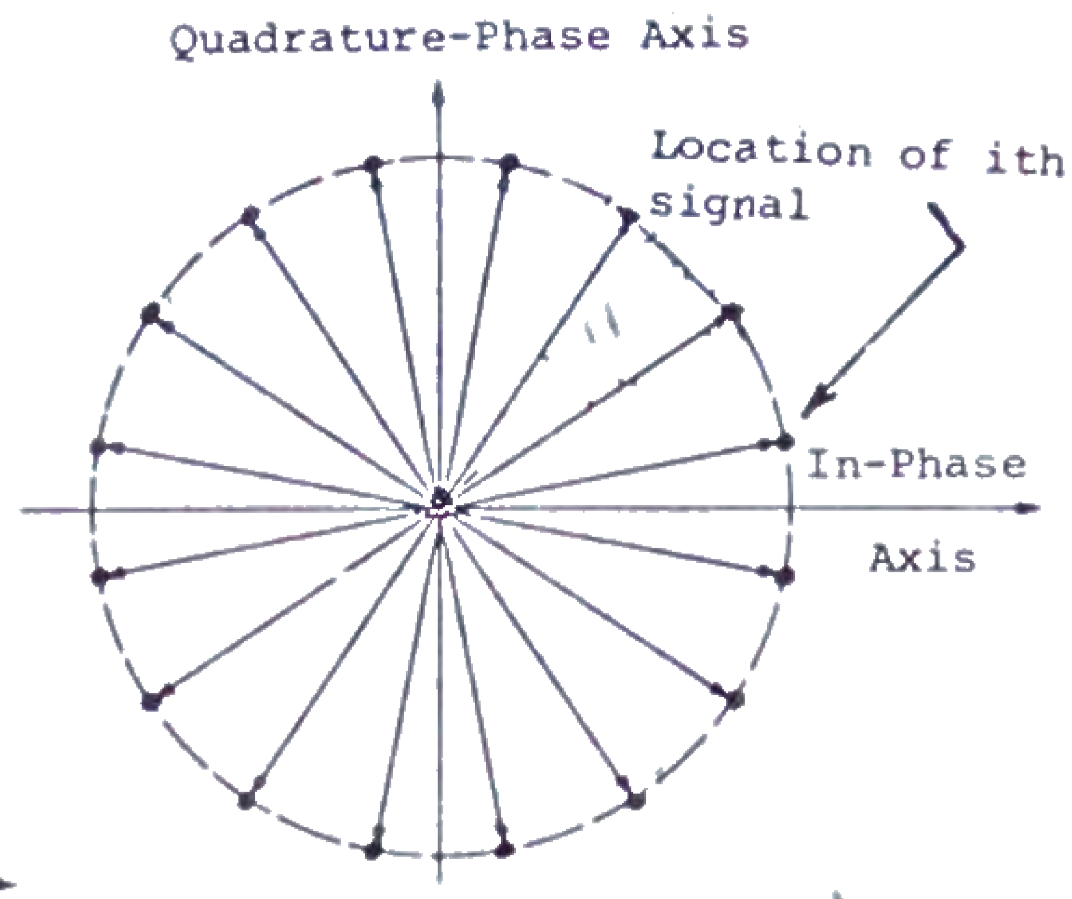


Figure 6a

16-PSK Signal space

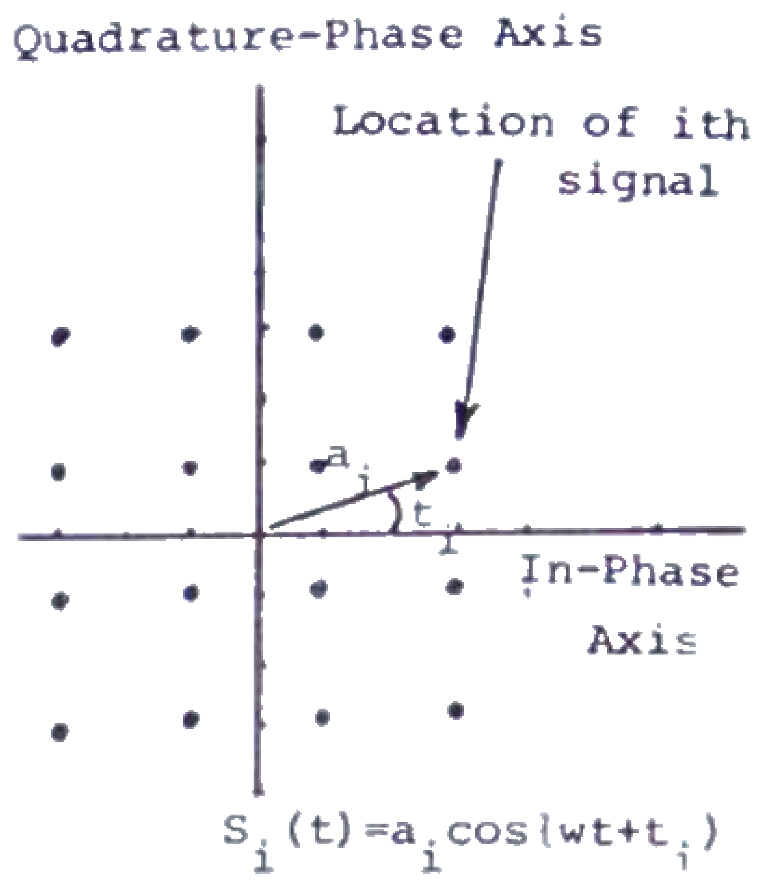


Figure 6b

16-QAM Signal space

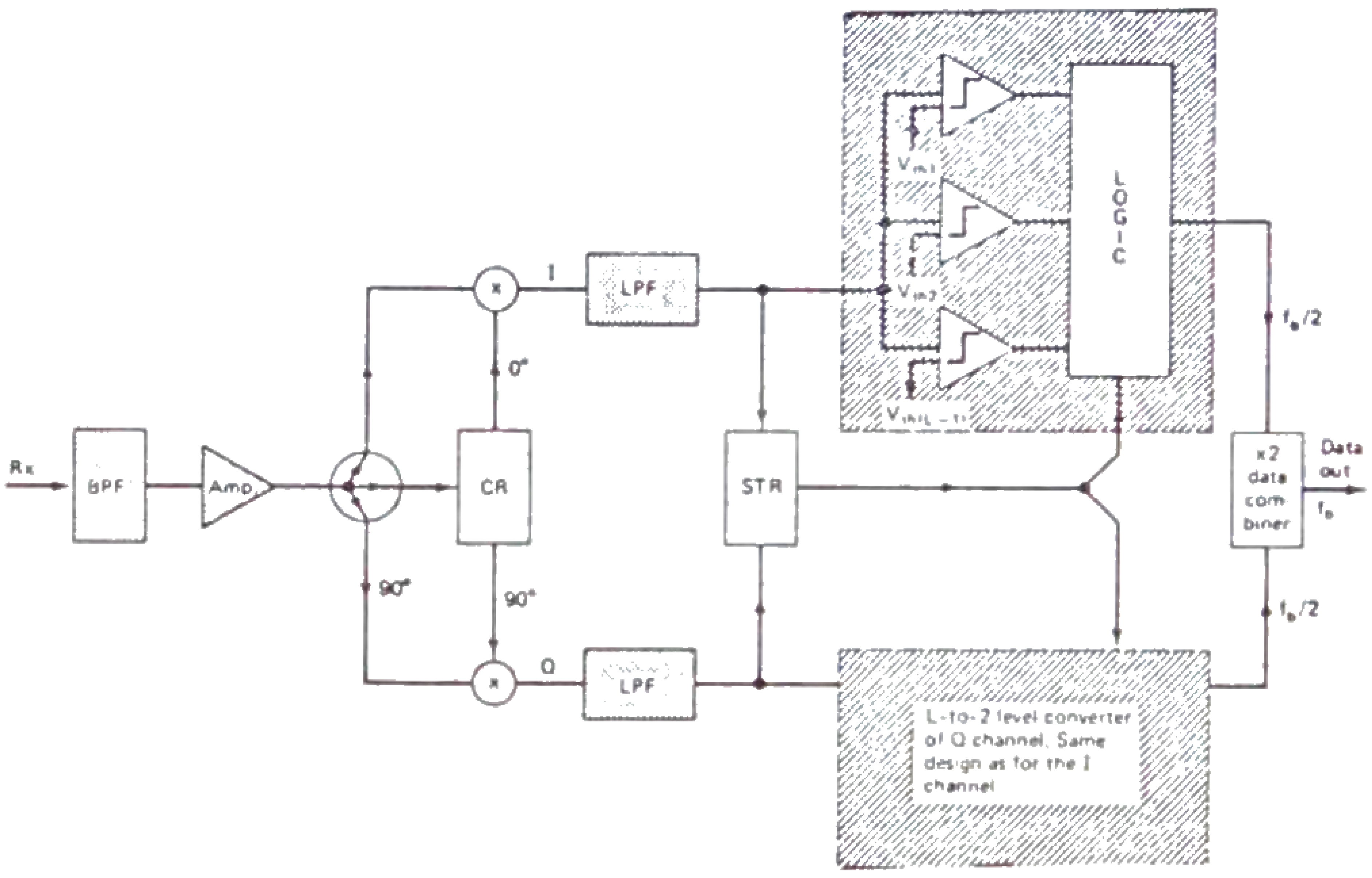


Figure 8: (Reference11, pg. 126, Fig. 6.10)

M-ary QAM or M-ary PSK demodulator block diagram

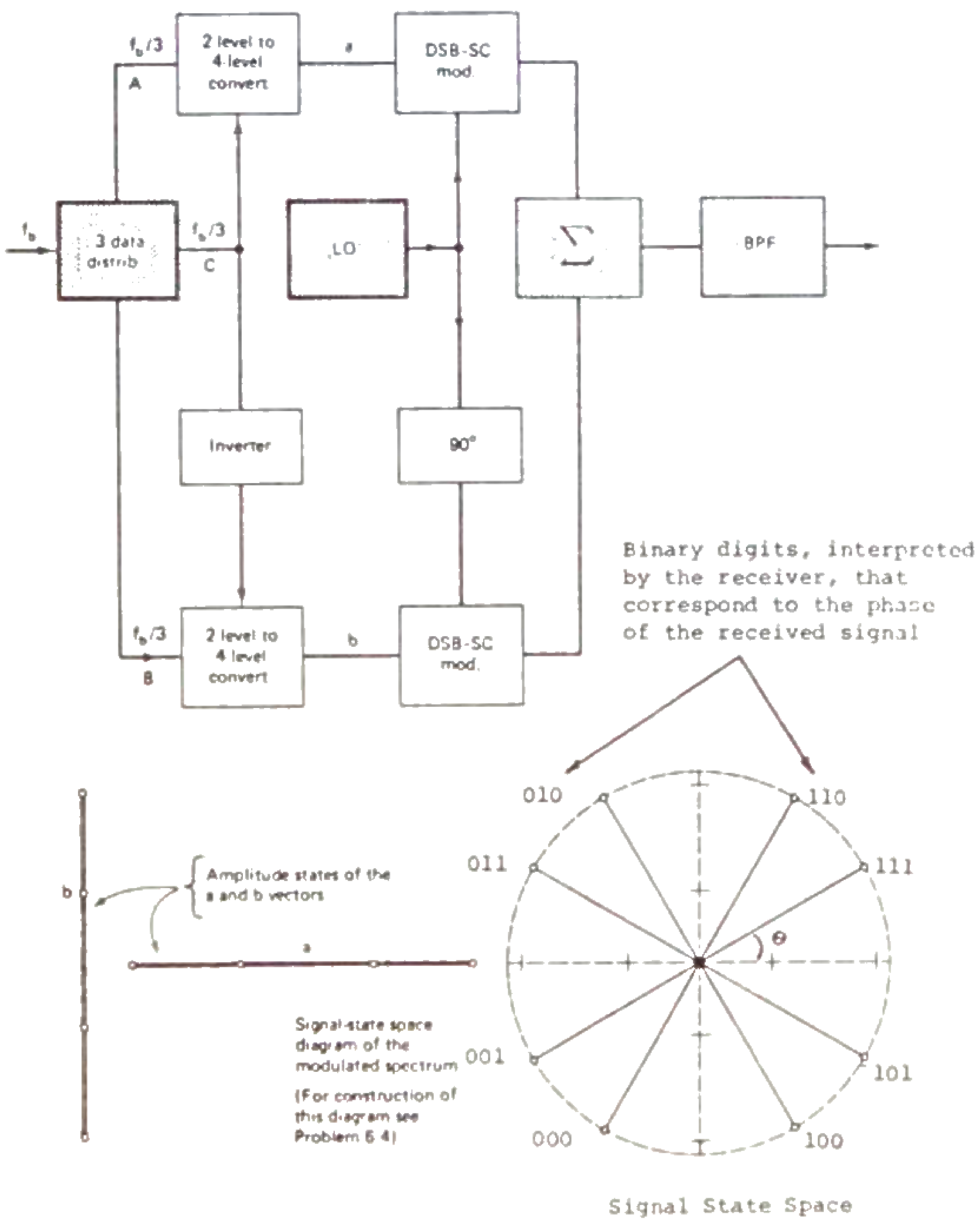


Figure 7: (Reference 11, pg. 120, Fig. 6.4)

8-PSK modulator block diagram and signal state space

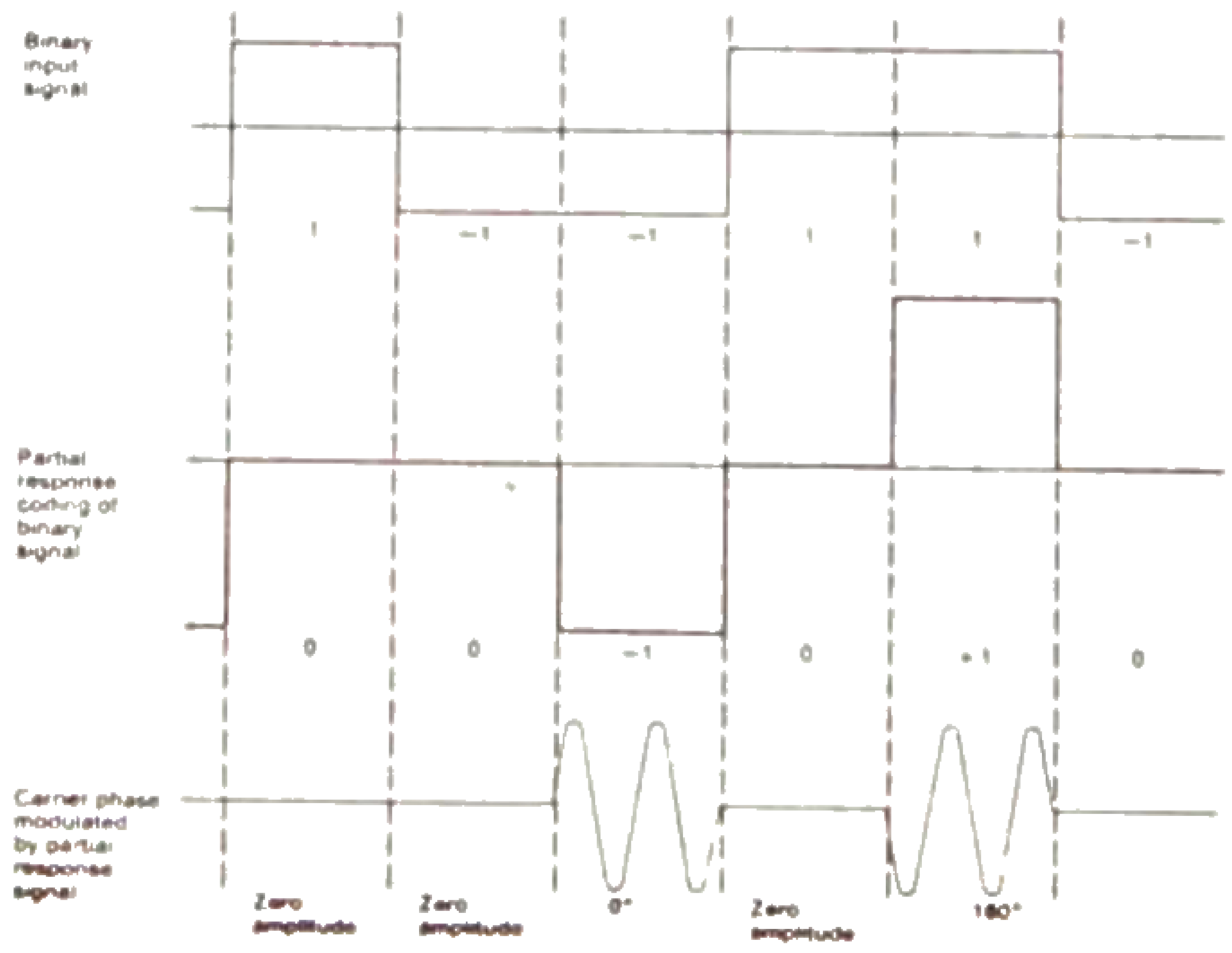


Figure 9: (Reference 2, pg 175, Fig. 3)

Partial Response Signalling:
The binary signal is modified so that the output signal for any symbol is the average of two consecutive input symbols. The original two-level signal has been converted into a three-level signal with a more suitable spectrum.

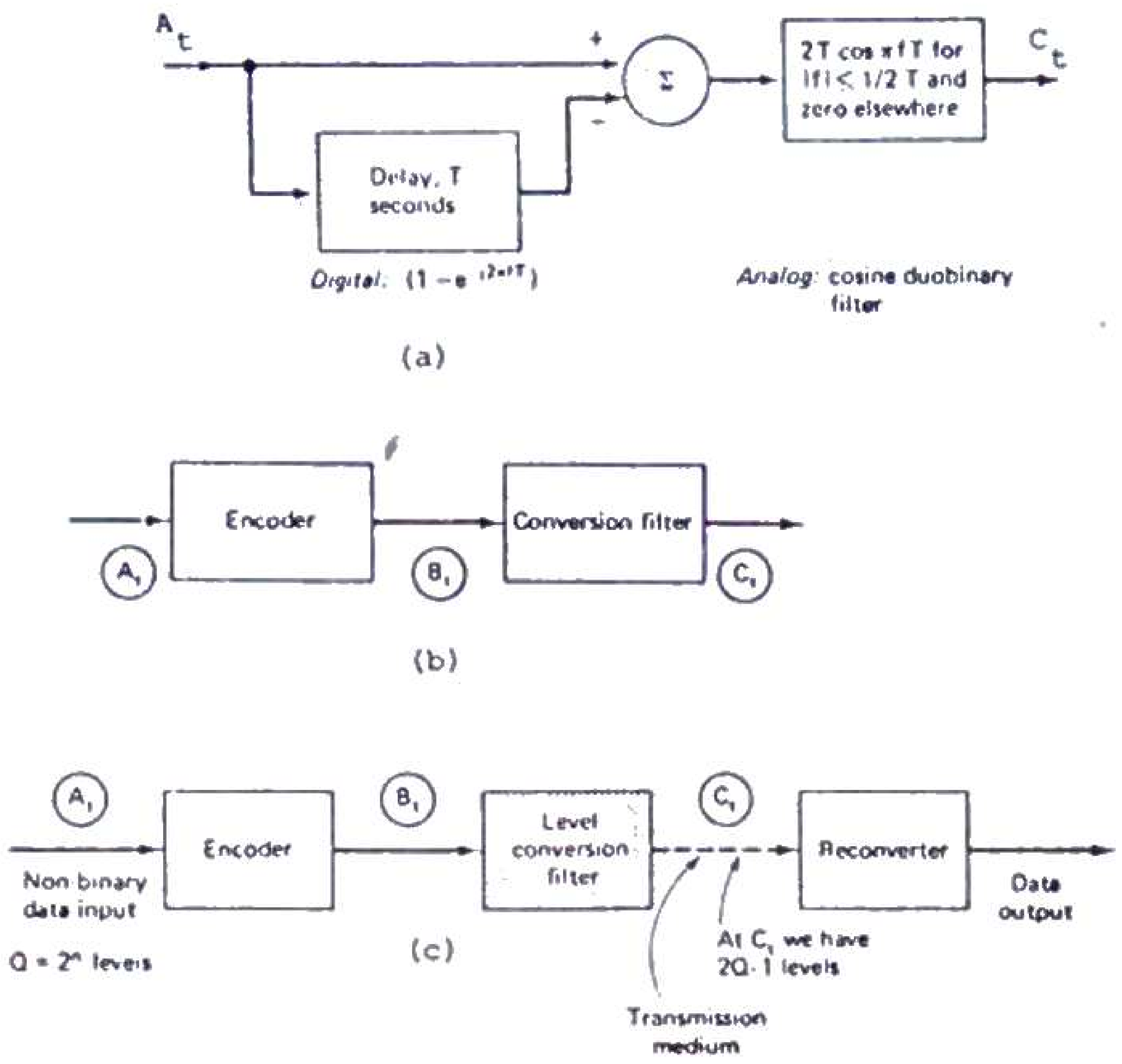


Figure 10: (Reference 11, pg 151,152,157)

Partial Response Signalling

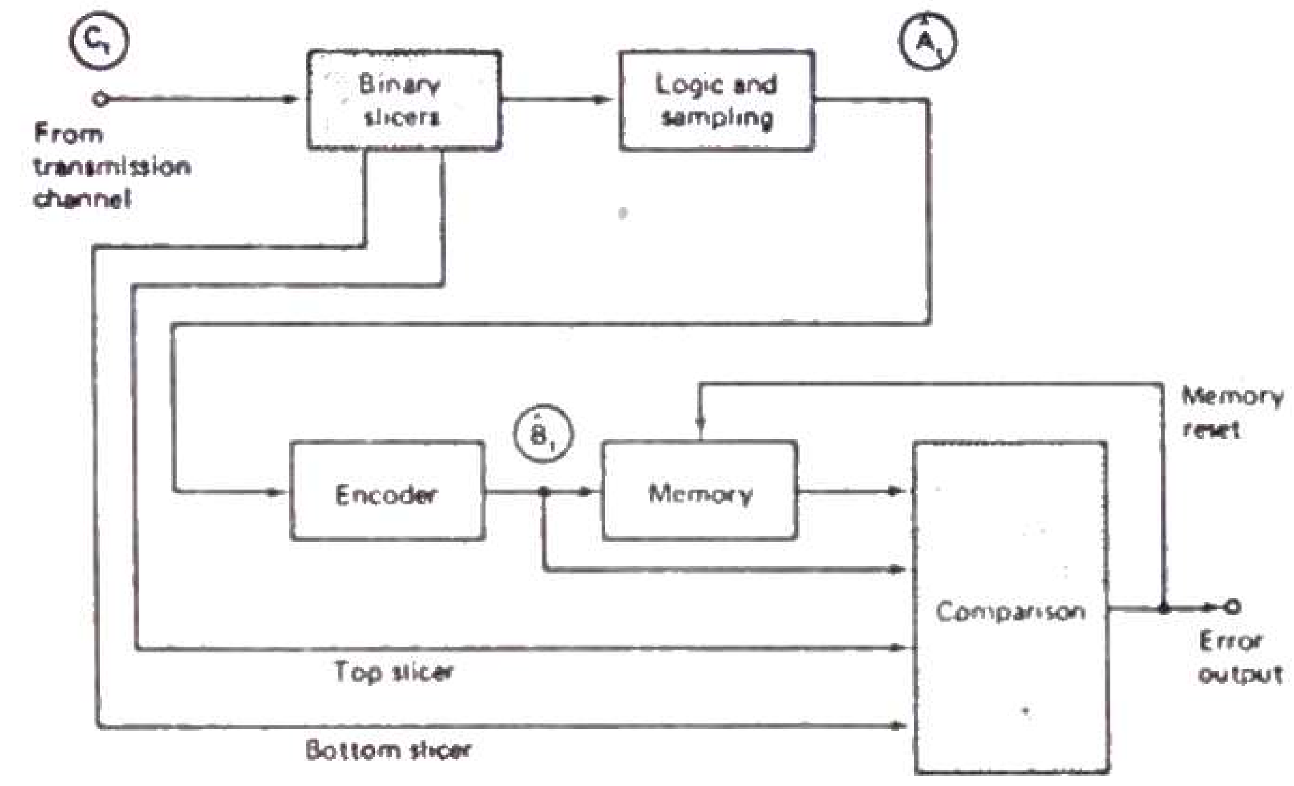


Figure 11:(Reference 11, pg. 164, Fig. 7.17)

Error detection for Correlative systems

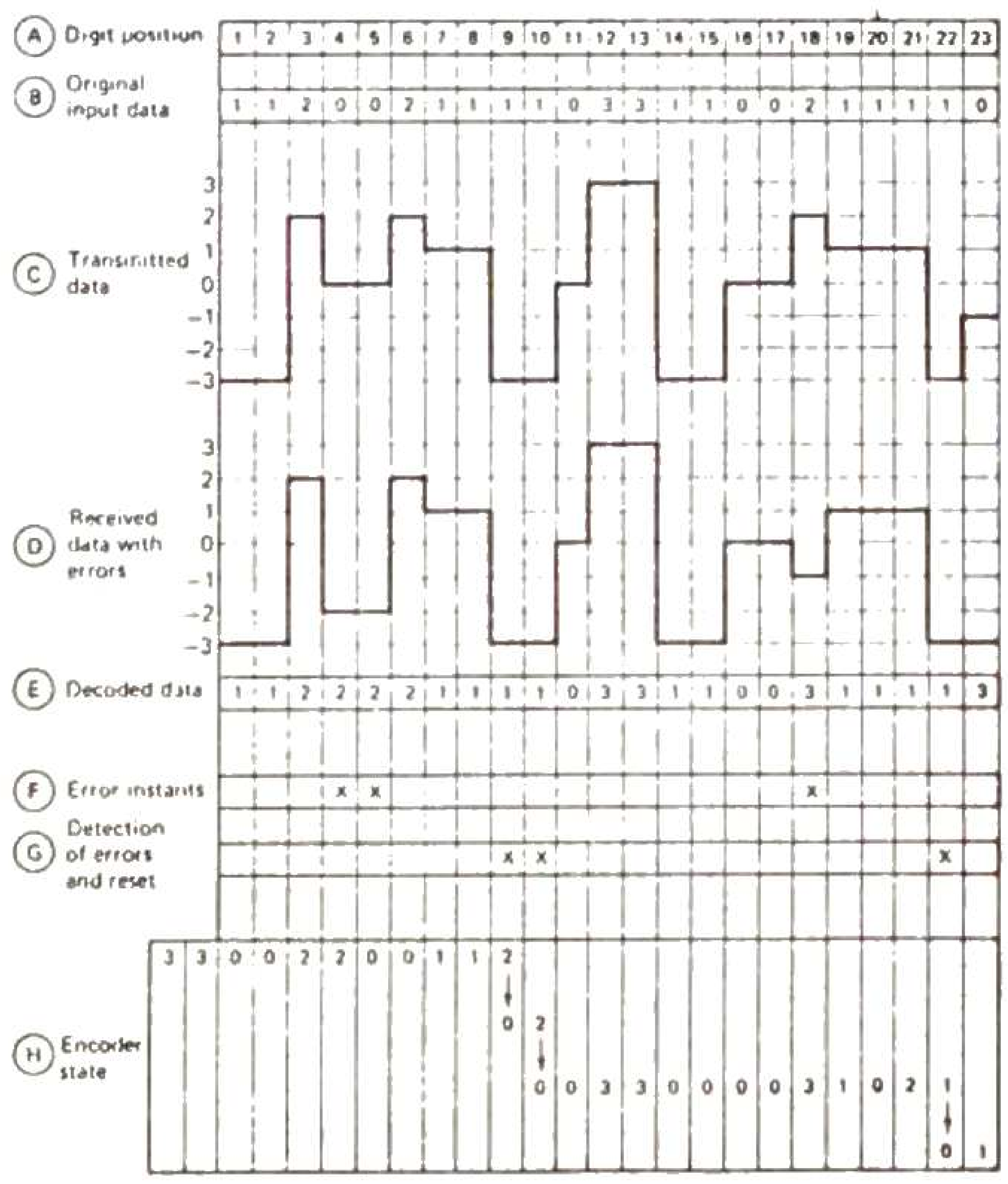


Figure 12:

(Reference11, pg. 167, Fig. 7.19)

Error detection patterns for Q=4

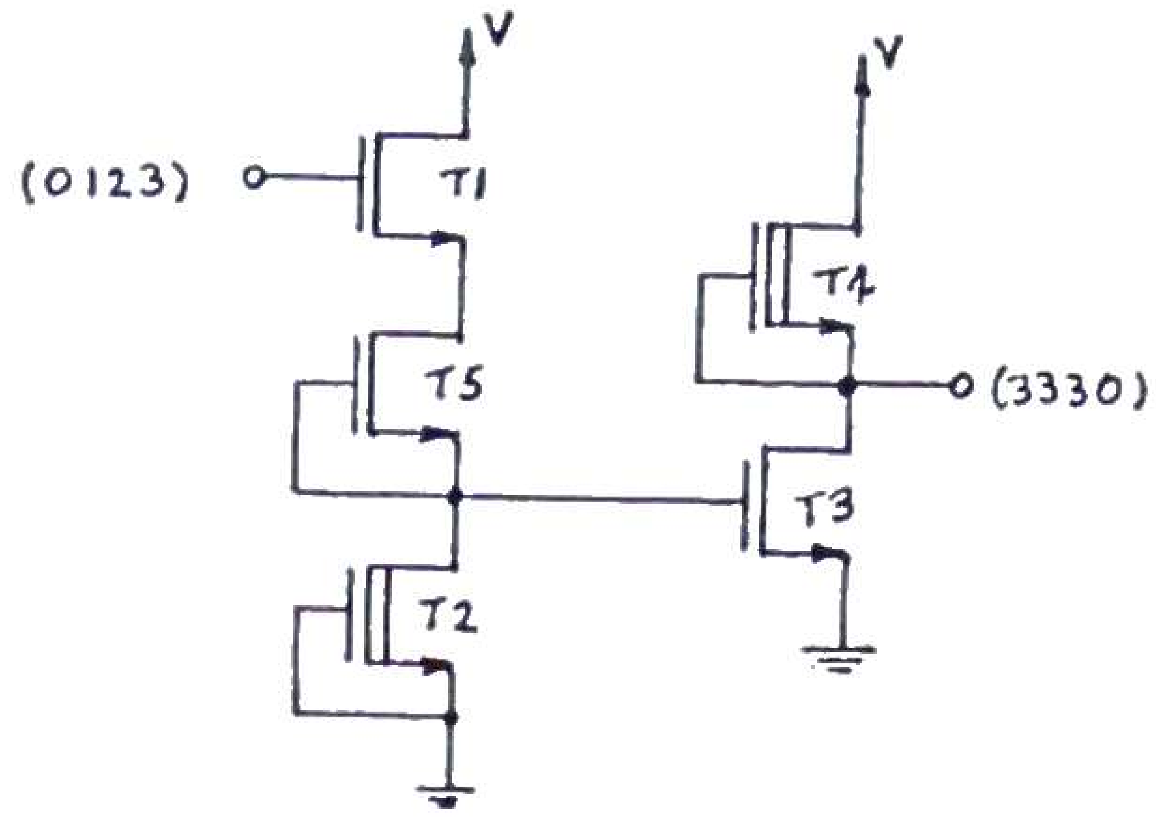


Figure 13: (Reference 21, pg. 628, Fig. 15)

A Quaternary NMOS High Threshold Detector

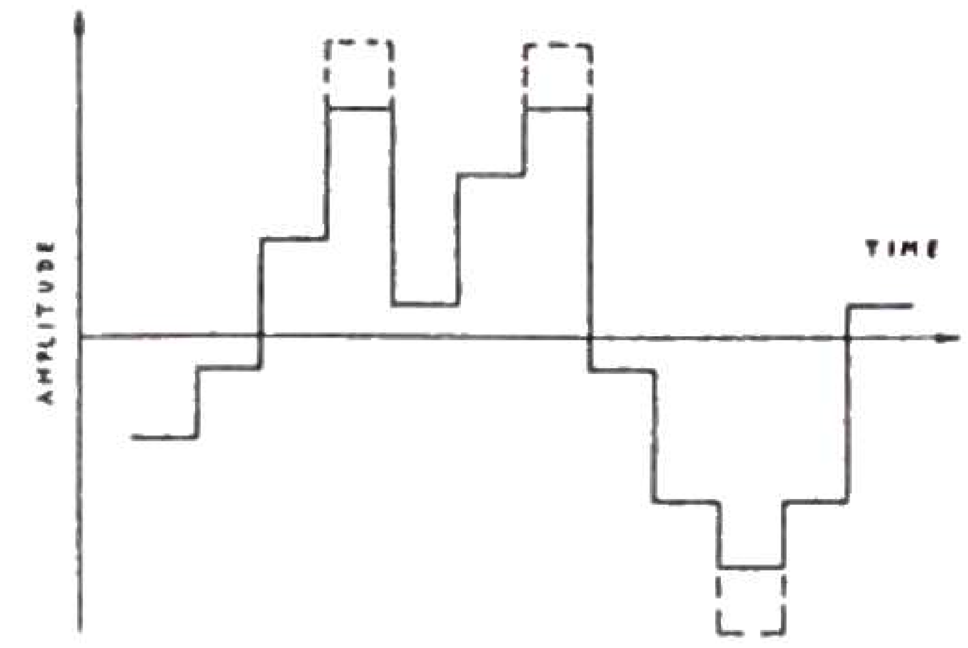
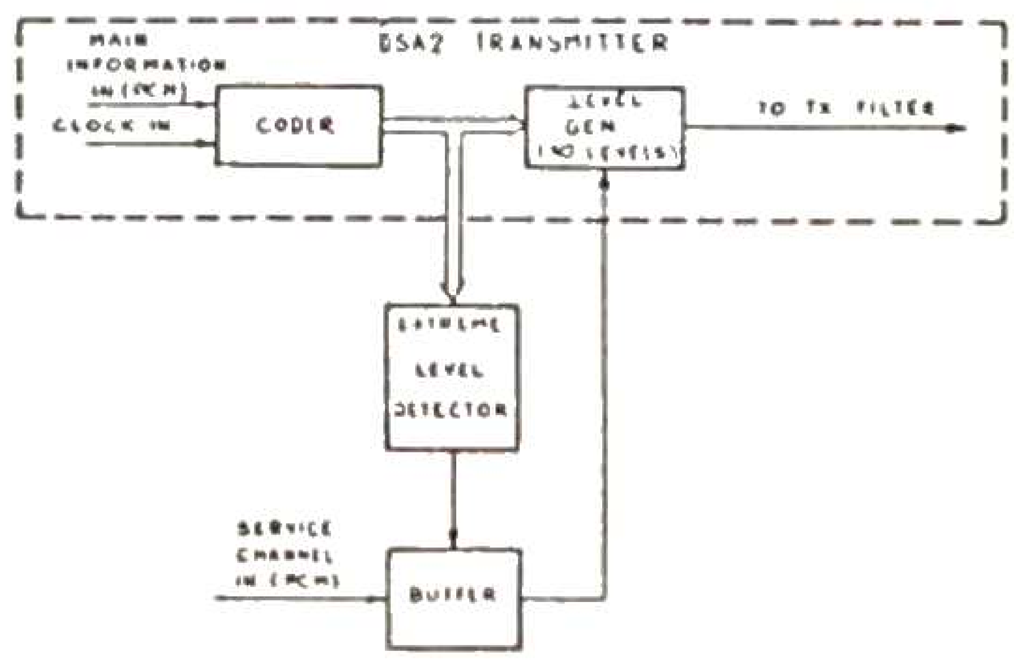
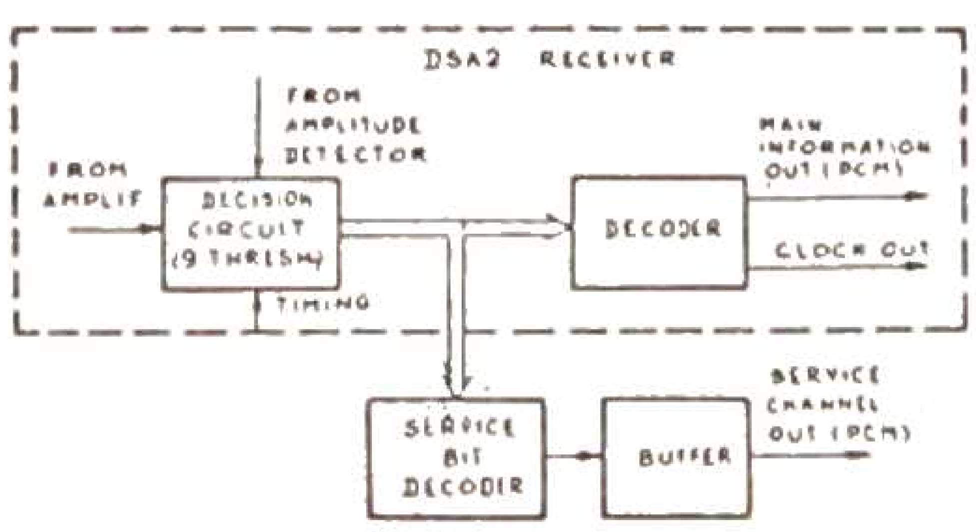


Figure 14: (Reference 13, pg. 1947, Fig 2, 3, 4)

14a: The dashed lines are the modifications of the PAM waveform corresponding to the transmission of a supplementary information bit



14b: Addition of service channel by LDM, transmitting end



14c: Recovery of service channel at receiving end